On convergence of higher order schemes for the projective integration method for stiff ordinary differential equations

John Maclean*, Georg A. Gottwald

School of Mathematics and Statistics, University of Sydney, NSW 2006. Australia.

Abstract

We present a convergence proof for higher order implementations of the projective integration method (PI) for a class of deterministic multi-scale systems in which fast variables quickly settle on a slow manifold. The error is shown to contain contributions associated with the length of the microsolver, the numerical accuracy of the macrosolver and the distance from the slow manifold caused by the combined effect of micro- and macrosolvers, respectively. We also provide stability conditions for the PI methods under which the fast variables will not diverge from the slow manifold. We corroborate our results by numerical simulations.

Keywords: multi-scale integrators, projective integration, error analysis
2010 MSC: 65LXX, 65PXX, 34E13, 37MXX

1. Introduction

Many problems in the natural sciences are modelled by multidimensional ordinary differential equations with entangled processes running on widely separated time scales. One is often interested in resolving the behaviour of the slow processes over a long, macro time scale. However, the fast processes prevent direct solution of the system by traditional numerical methods. Recently two numerical methods designed to overcome the restriction to the small integration time step in these stiff dynamical systems have been much studied; the *projective integration method* within the equation-free framework and the *heterogeneous multiscale methods* (HMM). Each method exists in multiple formulations; in the PI method, we mention [1, 2, 3, 4, 5, 6, 7, 8], and in the HMM, [9, 10, 11, 12, 13, 14, 15]. There is some debate on the similarities and differences between the methods; the interested reader is referred to [16, 17] for

---

*Corresponding author
Email addresses:* `j.maclean@maths.usyd.edu.au` (John Maclean),
`georg.gottwald@sydney.edu.au` (Georg A. Gottwald)

a discussion.

Both methods assume that the fast variables in the full multiscale system quickly relax to a slow manifold, after which the dynamics of the slow variables is governed by a slow reduced system. Both methods estimate the effective influence of the fast variables on the dynamics of the slow variables by employing a *microsolver* to perform short fine-scale computations with small time steps (microsteps). This information is used to propagate the dynamics on the slow manifold for large time steps (macrosteps) in the *macrosolver*.

The philosophy behind each method is slightly different. The PI approach estimates the effective slow vector field via direct numerical evaluation, not assuming any knowledge on the form of the reduced vector field; this forms part of the equation-free approach. In contrast, the HMM philosophy utilises a priori analytical knowledge about the reduced vector field.

In this paper, we focus on numerical methods that are *seamless*; that is, the numerical methods do not explicitly separate the slow variables and the fast variables at any stage in the solver, but instead propagate all variables simultaneously. These methods are useful in systems where conceptually there exists a decomposition or transformation of the system into slow and fast variables, but where this transformation is unknown. The added complication of seamless numerical methods is that the fast variables are propagated simultaneously with the slow variables with the large time step of the macrosolver. This may lead to a more severe departure of the fast variables from the slow manifold over the macrosteps in comparison to nonseamless methods.

In first order PI methods the micro- and macrosolver are applied sequentially, so the error accrued by the micro- and macrosolver can be analysed separately, as for example in [11, 18]. There are two different approaches to extend PI to higher order solvers. First, one can still apply the micro- and macrosolver sequentially, as in [19, 12, 5, 20]. The analysis in [12, 20] shows that such schemes can be accurate to second order in the size of the macrosolver. Alternatively, one can apply the microsolver multiple times during each time step of the macrosolver, as in [11, 4, 21]. The numerical schemes that we will consider take this approach. The analysis of such methods is complicated by the requirement that the errors accrued by the micro- and macrosolvers, which are intertwined due to the nonlinear nature of the dynamics, have to be estimated simultaneously. In [11], an error bound is proposed for a seamless HMM scheme of arbitrary order, albeit without proof. In [4, 5, 20], second order PI schemes are proposed and analysed. In [21], error bounds for the slow variables and stability conditions are derived for an arbitrary order Runge-Kutta macrosolver applied to a kinetic equation with linear relaxation.

In this paper we present a higher order seamless multiscale method as considered in [11, 4], for a system of nonlinear stiff ordinary differential equations. We propose a slight modification of this method which, involving an additional application of the microsolver, constructs slow vector fields pointing towards the

slow manifold. Both schemes reduce to Runge-Kutta methods if the microsolver is switched off. We establish rigorous convergence results for the slow variables of these methods. We find that both methods incur error terms propotional to the order of the macrosolver, the distance of the fast variables from the slow manifold, and an additional term due to the microsolver, independent of the order of the microsolver. This result confirms for the two methods we consider the error bound suggested in [11]. Furthermore, we find that the error due to the microsolver is smaller in our proposed method when both methods are employed at the same computational cost.

A known problem in seamless methods is that the macrosolver may lead to a departure of the fast variables from the slow manifold. To combat this divergence of the fast variables, several methods have been introduced [22, 23, 24, 25]; analytical bounds on the departure of the fast variables from the slow manifold over a macrostep have received relatively little attention (with the notable exception of [12]). Estimates of the maximal deviation of the fast variables from the slow manifold are particularly important when bifurcations occur or when the dynamics transits to different solution branches (e.g. [19, 1, 7, 26]); if the departure from the slow manifold is too large, the transitions may be premature.

We establish bounds on the departure of the fast variables from the slow manifold over the macrosolver. The bounds show that the numerically induced departure of the fast variables from the slow manifold scales one order better in the macrostep size in our modified version of PI. Furthermore, these bounds allow us to derive stability conditions for both methods under which the departure of the fast variables from the slow manifold remains finite over the macrosteps.

The paper is organized as follows. In Section 2 we discuss the class of dynamical systems studied, and briefly summarize in Section 3 classic Runge-Kutta methods for these systems. We then present two multiscale methods which enable the solution of these systems with macro length time steps in Section 4. In Section 5, the main part of this work, we derive rigorous error bounds for those numerical multiscale methods. In Section 6 we present results from numerical simulations corroborating our analytical findings. We conclude with a discussion in Section 7.

2. Model

We consider deterministic multiscale systems of the form

$$\dot{z}_\varepsilon = \mathcal{F}(z_\varepsilon, \varepsilon) \,, \tag{2.1}$$

with $z_\varepsilon \in \mathbb{R}^{n+m}$ and time scale separation parameter $0 < \varepsilon \ll 1$. We assume there is a (possibly unknown) decomposition $z_\varepsilon = (x_\varepsilon, y_\varepsilon)$ into fast variables $x_\varepsilon \in \mathbb{R}^m$ and slow variables $y_\varepsilon \in \mathbb{R}^n$ which evolve according to

$$\dot{y}_\varepsilon = g(x_\varepsilon, y_\varepsilon) \,, \tag{2.2}$$

$$\dot{x}_\varepsilon = \frac{1}{\varepsilon} f(x_\varepsilon, y_\varepsilon) . \tag{2.3}$$

We consider here the particular fast vector fields of the form

$$f(x_\varepsilon, y_\varepsilon) = \frac{\Lambda}{\varepsilon}(-x_\varepsilon + h_0(y_\varepsilon)) . \tag{2.4}$$

We assume there is a coordinate system such that the matrix $\Lambda \in \mathbb{R}^{m \times m}$ is diagonal with diagonal entries $\lambda_{ii} > 0$. We further allow for a scaling of time such that $\min(\lambda_{ii}) = 1$ and define $\max(\lambda_{ii}) = \lambda$. We assume that there exists a slow manifold $x = h_\varepsilon(y) = h_0(y) + \mathcal{O}(\varepsilon)$, towards which initial conditions are attracted exponentially fast. On the slow manifold, the dynamics slows down and is approximately determined by

$$\dot{Y} = G(Y) , \tag{2.5}$$

with $Y = y_\varepsilon + \mathcal{O}(\varepsilon)$ and reduced slow vectorfield

$$G(y) = g(h_\varepsilon(y), y) . \tag{2.6}$$

3. Runge-Kutta Solvers

We denote by $z_\varepsilon(t^n)$ the solution of (2.1) evaluated at the discrete time $t^n = n\Delta t$, and by $\bar{z}^n$ the numerical approximation of $z_\varepsilon(t^n)$ given by a Runge-Kutta solver of order P. Runge-Kutta solvers form approximations to the dynamics in terms of *increments*. For simplicity, we restrict our analysis to Runge-Kutta methods in which increments are given recursively by

$$\bar{k}_j(\bar{z}^n) = \Delta t \mathcal{F}(\bar{z}^n + a_j \bar{k}_{j-1}, \varepsilon) , \tag{3.1}$$

for $j = 1, 2, \ldots, P$. The values of the nodes a_j depend on the order P (see for instance [27]), and satisfy $0 \le a_j \le 1$, with $a_1 = 0$ so that the first increment is defined explicitly. Each increment $\bar{k}_j$ evaluates the vector field $\mathcal{F}$ of (2.1) at the intermediate time $t^n + a_j \Delta t$. The increments are averaged to define $\bar{z}^{n+1}$, with

$$\bar{z}^{n+1} = \bar{z}^n + \sum_{j=1}^{P} b_j \bar{k}_j(\bar{z}^n) , \tag{3.2}$$

where the weights b_j satisfy the condition $\sum_{j=1}^{P} b_j = 1$, and depend on the order P and the particular choice of nodes a_j. For instance, for $P = 4$, the widely used fourth-order Runge-Kutta scheme, the nodes and weights may be given by

$$a_j = \left\{ 0, \frac{1}{2}, \frac{1}{2}, 1 \right\} , \tag{3.3}$$

$$b_j = \left\{ \frac{1}{6}, \frac{1}{3}, \frac{1}{3}, \frac{1}{6} \right\} . \tag{3.4}$$

For any P, the nodes and weights are determined such that the application of a single Runge-Kutta step of order P to a system with initial condition $\bar{z}_\varepsilon(t)$ and time step Δt produces an approximation to $\bar{z}_\varepsilon(t+\Delta t)$ accurate to within $\mathcal{O}(\Delta t^{P+1})$; see for instance [27]. In particular for linear systems

$$\dot{x}_\varepsilon = -\frac{\Lambda}{\varepsilon} x_\varepsilon \ ,$$

for which

$$\frac{\mathrm{d}^j x_\varepsilon}{\mathrm{d}t^j} = \left(-\frac{\Lambda}{\varepsilon}\right)^j \ ,$$

a single Runge-Kutta step of P-th order can be written as

$$x^{n+1} = \rho\left(-\frac{\Lambda \Delta t}{\varepsilon}\right) x^n \ ,$$

where the *linear amplification factor* ρ is given by the Taylor polynomial to order P of an exponential function

$$\rho(\eta) = \sum_{j=0}^{P} \frac{(\eta)^j}{j!} \ . \tag{3.5}$$

A straightforward implementation of Runge-Kutta methods to simulate stiff dynamical systems such as (2.2)–(2.4) would be computationally too costly, as the time step is restricted to $\Delta t \leq \mathcal{O}(\varepsilon)$ to ensure numerical stability.

In the next section we present two numerical multiscale schemes which are designed to overcome the problem of stiffness presented above. These schemes employ a microsolver to relax the fast variables towards the slow manifold. Utilising the slowness of the dynamics on the slow manifold allows for the application of Runge-Kutta methods with large macro time steps $\Delta t \gg \varepsilon$.

4. Numerical Multiscale Methods

We consider two seamless projective integration methods. The first is a general order formulation of PI as proposed in [11, 4, 12]. We call this scheme PI1. The second is a modification of PI1, which employs information from the microsolver to define increments which point in the direction of the slow manifold, at the cost of one additional application of the microsolver[1]. We call this method PI2. The PI1 and PI2 schemes differ in the definition of the increments.

---

[1]We ensure that the overall cost of PI1 and PI2 is the same when they are compared numerically by adjusting the total number of microsteps in each method (see Section 6).

Denote by z^n the numerical approximation given by the multiscale scheme to $z_\varepsilon(t^n)$; using z^n as the initial condition, both methods employ a microsolver with small microstep δt, and then evaluate the vectorfield over a large macrostep $\Delta t \gg \varepsilon$. Iterating these steps enables one to construct increments which cover a macro time scale. The macrosolver then combines these increments in a weighted sum in Runge-Kutta fashion.

We denote by $\varphi^{m,\delta t}$ the flow map for the microsolver run for m microsteps with time step δt and assume that it describes an explicit numerical method of order p. We do not specify which particular numerical method is chosen; as we will see in Proposition 5.6, increasing the order of the microsolver does not improve the predicted overall error scaling.

In the following we detail PI1 and PI2 and highlight their differences. The procedures are illustrated in Figure 1 for PI1 and in Figure 2 for PI2.

4.1. Projective Integration Scheme PI1

We describe here a general order formulation of projective integration along the lines of [11, 4, 12, 18]. We remark that this formulation is an instance where PI and HMM are essentially the same (see [11, 16]). The scheme PI1 is a modified Runge-Kutta scheme in which the microsolver is employed to relax the fast variables close to the slow manifold before each increment is estimated.

We denote by $z_m^{n,j}$ the approximation of the fast and slow variables at the m-th microstep of the j-th increment at time step n, and denote by M_j the integer number of microsteps taken before the j-th increment is estimated. We denote discrete times associated with microsolvers by subscripts and those associated with macrosolvers by superscripts.

The increments cover a time step of Δt and are given by evaluating $\mathcal{F}$ after an application of the microsolver, with

$$\hat{k}_j(z^n) = \Delta t\,\mathcal{F}(z_{M_j}^{n,j}, \varepsilon)\,, \tag{4.1}$$

for $j = 1, 2, \ldots, P$, where we define $z_m^{n,j}$ for $j = 1, 2, \ldots, P$, $m = 1, 2, \ldots, M_j$, as the output of the microsolver

$$z_m^{n,j} = \varphi^{m,\delta t}\left(z_0^{n,j}\right)\,, \tag{4.2}$$

with initial condition

$$z_0^{n,j} = \begin{cases} z^n & \text{for } j = 1 \\ z_{M_1}^{n,1} + a_j \hat{k}_{j-1}(z^n) & \text{for } j > 1 \end{cases}\,. \tag{4.3}$$

The nodes a_j are those used in the increments of a Runge-Kutta solver of order P; i.e for $P = 4$, a_j may be given by (3.3). For more general Runge-Kutta solvers for PI methods, see [21]. Construction of the microsteps $z_m^{n,j}$ is illustrated in Figures 1a and 1c, and construction of the increments $\hat{k}_j$ in Figures 1b and 1d.

The macrosolver is then given by the weighted sum

$$z^{n+1} = z_{M_1}^{n,1} + \sum_{j=1}^{P} b_j \hat{k}_j(z^n) \ , \qquad (4.4)$$

where the weights b_j are appropriate to a Runge-Kutta solver of order P; i.e for $P = 4$, b_j may be given by (3.4). The macrosolver is illustrated in Figure 1e. Note that for $M_j = 0$ for all j, i.e. without the microsolver, the scheme reduces to a standard Runge-Kutta solver of order P applied to the system (2.1). It is not true that PI schemes in general reduce to a numerical discretisation of the underlying multi-scale dynamical system if the microsolver is switched off (see for example [20]).

For the analysis of the PI1 scheme, it is helpful to explicitly identify the slow and fast variables. We therefore decompose the PI1 variables z^n into fast and slow components (x^n, y^n), and $z_{M_j}^{n,j}$ into the fast and slow components $(x_{M_j}^{n,j}, y_{M_j}^{n,j})$. Furthermore, we split the PI1 increments $\hat{k}_j$ into fast components $\hat{k}_{x,j}$ and slow components $\hat{k}_{y,j}$, with

$$\hat{k}_{y,j}(x^n, y^n) = \Delta t \, g\left(x_{M_j}^{n,j}, y_{M_j}^{n,j}\right) \ , \qquad (4.5)$$

$$\hat{k}_{x,j}(x^n, y^n) = \Delta t \, f\left(x_{M_j}^{n,j}, y_{M_j}^{n,j}, \varepsilon\right) \ . \qquad (4.6)$$

The macrosolver is then written as

$$y^{n+1} = y_{M_1}^{n,1} + \sum_{j=1}^{P} b_j \hat{k}_{y,j}(x^n, y^n) \ ,$$

$$x^{n+1} = x_{M_1}^{n,1} + \sum_{j=1}^{P} b_j \hat{k}_{x,j}(x^n, y^n) \ . \qquad (4.7)$$

4.2. Projective Integration Scheme PI2

We present here a modification of the PI1 scheme in which the increments are given by differences between endpoints of the microsolver. We again denote by $z_m^{n,j}$ the approximation of the fast and slow variables at the m-th microstep of the j-th increment. The PI1 increments are given by (4.1), which we recall here as

$$\hat{k}_j(z^n) = \Delta t \, \mathcal{F}(z_{M_j}^{n,j}, \varepsilon) \ , \qquad (4.8)$$

for $j = 1, 2, \ldots, P$, where $z_m^{n,j}$ is now defined for $j = 1, 2, \ldots, P+1$, $m = 1, 2, \ldots, M_j$, as the output of the microsolver

$$z_m^{n,j} = \varphi^{m,\delta t}\left(z_0^{n,j}\right) \ , \qquad (4.9)$$

with initial condition

$$z_0^{n,j} = \begin{cases} z^n & \text{for } j=1 \\ z_{M_1}^{n,1} + a_j \hat{k}_{j-1}(z^n) & \text{for } j>1 \end{cases} .$$ (4.10)

Construction of the microsteps $z_m^{n,j}$ is illustrated in Figures 2a, 2c and 2e. The PI2 increments are constructed by approximating the vector field $\mathcal{F}$ according to $\mathcal{F}(z^{n_j,M_j}) \approx (z_{M_{j+1}}^{n,j+1} - z_{M_1}^{n,1})/(a_{j+1}\Delta t)$, leading to

$$k_j(z^n) = \frac{1}{a_{j+1}} \left(z_{M_{j+1}}^{n,j+1} - z_{M_1}^{n,1} \right) ,$$ (4.11)

for $j=1,2,\ldots,P$. The nodes a_j with $j=1,2,\ldots,P$ are again those used in the increments of a Runge-Kutta solver of order P, and we set $a_{P+1}=1$. The construction of the PI2 increments k_j is illustrated in Figures 2c and 2e.

Each PI2 increment covers a time step of $\Delta t + M_{j+1}\delta t/a_{j+1}$. We fix the total number of microsteps M_j for $j>1$ with

$$M_j = a_j M ,$$ (4.12)

for $j=1,2,\ldots,P+1$ and for some M satisfying $a_j M \in \mathbb{N}$, so that each increment covers a uniform time step of $\Delta t + M\delta t =: t_\Delta$. Note that (4.12) allocates more microsteps after larger increments and less after shorter increments.

The macrosolver is now constructed as a weighted sum over the relaxed increments k_j rather than over $\hat{k}_j$, with

$$z^{n+1} = z_{M_1}^{n,1} + \sum_{j=1}^{P} b_j k_j(z^n) ,$$ (4.13)

where the weights b_j again correspond to a Runge-Kutta solver of order P. Note that $t^n = n(t_\Delta + M_1\delta t)$ for PI2. The macrosolver is illustrated in Figure 2f.

Again for $M_1 = M = 0$, i.e. without the microsolver, the PI2 scheme reduces to a standard Runge-Kutta method of order P.

As with the PI1 scheme, it is helpful to explicitly identify the slow and fast variables in the solver. We therefore decompose the PI2 variables z^n into fast and slow components (x^n, y^n), and $z_{M_j}^{n,j}$ into the fast and slow components $(x_{M_j}^{n,j}, y_{M_j}^{n,j})$, and we split the PI2 increments k_j into fast components $k_{x,j}$ and slow components $k_{y,j}$, with

$$k_{y,j}(x^n, y^n) = \frac{1}{a_{j+1}} \left(y_{M_{j+1}}^{n,j+1} - y_{M_1}^{n,1} \right) ,$$ (4.14)

$$k_{x,j}(x^n, y^n) = \frac{1}{a_{j+1}} \left(x_{M_{j+1}}^{n,j+1} - x_{M_1}^{n,1} \right) ,$$ (4.15)

depending via (4.10) on the PI1 increments $\hat{k}_j$. For completeness we recall these as

$$\hat{k}_{y,j}(x^n, y^n) = \Delta t\, g\left(x_{M_j}^{n,j}, y_{M_j}^{n,j} \right) ,$$ (4.16)

$$\hat{k}_{x,j}(x^n,y^n) = \Delta t\, f\left(x_{M_j}^{n,j}, y_{M_j}^{n,j}, \varepsilon\right)\,. \tag{4.17}$$

The macrosolver is then given by

$$y^{n+1} = y_{M_1}^{n,1} + \sum_{j=1}^{P} b_j k_{y,j}(x^n,y^n)$$

$$x^{n+1} = x_{M_1}^{n,1} + \sum_{j=1}^{P} b_j k_{x,j}(x^n,y^n)\,. \tag{4.18}$$

For ease of exposition we write the slow dynamics as

$$y^{n+1} = y^n + \tilde{g}(x^n,y^n)\,,$$

where the vector field of the slow variables in the PI2 macrosolver, $\tilde{g}$, is given by

$$\tilde{g}(x^n,y^n) = y_{M_1}^{n,1} - y^n + \sum_{j=1}^{P} b_j k_{y,j}(x^n,y^n)\,. \tag{4.19}$$

Comparing Figures 1 and 2, in the PI2 method the increments point in the approximate direction of the slow manifold, so that the macrostep initialises the fast variables close to the slow manifold after a macrostep. By comparison, the PI1 increments can depart from the slow manifold with larger scale separations or for initial conditions off the slow manifold.

5. Error analysis for Projective Integration

We provide rigorous error bounds for the slow variables of PI in the formulations PI1 (4.1)–(4.4) and PI2 (4.9)–(4.13), following the general line of proof used in [11]. Therein the result for PI1 was stated, albeit without explicit proof. Furthermore, we establish bounds on the departure of the fast variables from the slow manifold over the macrosolver, yielding stability conditions for the fast variables.

Throughout this work we assume the following conditions on the growth and smoothness of solutions of our system and on the numerical discretization parameters of PI.

Assumptions

A1: The zeroth order approximation of the slow manifold $h_0(y)$ is Lipschitz continuous; that is there exists a constant L_h such that

$$|h_0(y_1) - h_0(y_2)| \le L_h |y_1 - y_2|\,.$$

A2: The vectorfield $g(x,y)$ is Lipschitz continuous; that is there exists a constant L_g such that

$$|g(x_1,y_1) - g(x_2,y_2)| \leq L_g(|x_1 - x_2| + |y_1 - y_2|).$$

A3: The second order derivatives of h_0 are all bounded; that is there exists a constant $L_{h'}$ such that

$$\sup\left| \sum_{|\alpha|=2} \partial^\alpha h_0(y)\right| \leq L_{h'},$$

where we used multi-index notation.

A4: The vectorfield $g(x,y)$ is bounded for all x,y; that is there exists a constant C_g such that

$$C_g = \sup|g(x,y)|.$$

A5: The reduced slow dynamics $Y(t)$ is of class $C^{\max(P,p)}$; that is there exist constants C_P^* and C_p^* such that

$$C_P^* = \sup\left|\frac{\mathrm{d}^P Y(t)}{\mathrm{d}t^P}\right|,$$

$$C_p^* = \sup\left|\frac{\mathrm{d}^p Y(t)}{\mathrm{d}t^p}\right|,$$

and in particular there exists a constant C_2^* satisfying

$$C_2^* = \sup|\ddot{Y}(t)|.$$

A6: The total time Δt of the macrostep is sufficiently short so that, employing the practical constraint $M\delta t \leq \Delta t$,

$$L_G M\delta t \leq L_G \Delta t < \frac{1}{2}.$$

Remark 5.1. *Assumptions (A1)–(A2) imply that the reduced slow dynamics (2.5) is also Lipschitz continuous and there exists a constant $L_G \leq L_g(1 + L_h)$ such that*

$$|G(Y_1) - G(Y_2)| \leq L_G|Y_1 - Y_2|.$$

Assumption (A4) implies that the reduced slow dynamics is also bounded and there exists a constant $C_G \leq C_g$ such that

$$C_G = \sup|G(Y)|.$$

The global Lipschitz conditions can be relaxed to local Lipschitz conditions by the usual means.

We will establish bounds for the error $\mathcal{E}^n$ between the PI1 and PI2 estimate y^n and the solution of the full system $y_\varepsilon(t^n)$,

$$\mathcal{E}^n = |y_\varepsilon(t^n) - y^n|.$$

Theorem 5.2 (Convergence). *Consider schemes PI1 and PI2 run with a Runge-Kutta method of order P for the macrosolver and an explicit scheme of order p for the microsolver. Given assumptions (A1)–(A6), there exists a constant C such that on a fixed time interval T, for each n such that $nt_\Delta \leq T$, the error between the PI1 and PI2 estimates and the exact solution of the full multiscale system (2.1) are bounded by*

$$\mathcal{E}^n \leq C\left(t_\Delta^P + M\delta t + \varepsilon + \left(\frac{\varepsilon}{t_\Delta} + \rho^{aM}\left(-\frac{\delta t}{\varepsilon}\right)\right)|d_{\max}^n|\right).$$

Here $a = \min_{j>1} a_j$, ρ is the linear amplification factor (3.5) for the microsolver of order p measuring the attraction of the fast variables to the slow manifold over a microstep, and $|d_{\max}^n| := \max_{\substack{0 \leq i < n, \\ 1 \leq k \leq P+1}} |x_0^{i,k} - h_0(y_0^{i,k})|$ is the maximal deviation of the fast variables from the approximate slow manifold accrued over the integration time.

It is worthwhile to briefly discuss the bound on $\mathcal{E}^n$. The term proportional to t_Δ^P reflects the convergence of the underlying Runge-Kutta numerical scheme of order P in the macrosolver. The term proportional to $M\delta t$ is incurred by the drift of the slow variables over the microsteps before estimating the increments (regardless of the order p of the microsolver). The terms proportional to the time scale parameter ε represent the error made by the reduction as well as an additional error incurred during the drift of the slow variable over the microsteps. The term proportional to $(\varepsilon/t_\Delta + \rho^{aM}(-\delta t/\varepsilon))|d_{\max}^n|$ measures the mismatch between the slow vector field $g(x,y)$ after an application of the microsolver and the reduced vector field $G(y)$.

We also provide bounds on the deviation $|d^n|$ of the fast variables x^n from the slow manifold $h_0(y^n)$ for PI1 and PI2,

$$|d^n| = |x^n - h_0(y^n)|.$$

Theorem 5.3 (Stability of the fast variables). *Consider schemes PI1 and PI2 run with a Runge-Kutta method of order P for the macrosolver and a forward Euler scheme for the microsolver. Given assumptions (A1), (A3) and*

(A4), the fast variables do not diverge over the macrosolver, so that the largest deviation of the fast variables from the slow manifold $|d_{\max}^n|$ is finite, if

$$\frac{\lambda \Delta t}{\varepsilon}\left(1-\frac{\delta t}{\varepsilon}\right)^{aM}<1 \ .$$

Then the distance of the fast variables from the slow manifold after the n-th macrostep satisfies for the PI1 scheme the recurrence relation

$$|d^{n+1}|\leq\sum_{j=1}^{P}b_j\left(\frac{\lambda \Delta t}{\varepsilon}\left(1-\frac{\delta t}{\varepsilon}\right)^{aM}\right)^{j}|d^n|+L_h C_g(1+\lambda)\Delta t \ ,$$

and for the PI2 scheme

$$|d^{n+1}|\leq\frac{\left(1-\frac{\delta t}{\varepsilon}\right)^{aM}}{a}\sum_{j=1}^{P}b_j\left(\frac{\lambda \Delta t}{\varepsilon}\left(1-\frac{\delta t}{\varepsilon}\right)^{aM}\right)^{j}|d^n|+2L_{h'}C_g^2 t_\Delta^2 \ .$$

Remark 5.4. *The stability condition for the fast variables of the PI1 and PI2 methods is identical to the corresponding stability condition for an Euler macrosolver given by Assumption 8 in [18] (see also [21]).*

We note that Theorem 5.3 can be formulated for a microsolver of order $p>1$ but, as we shall see, optimal convergence results are given by a forward Euler microsolver.

We briefly discuss the stability condition and the bounds for $|d^{n+1}|$ established above. The stability condition can be understood as follows: $(1-\delta t/\varepsilon)^{aM}$ denotes the exponential contraction of the fast variables towards the slow manifold during the application of the microsolver; if this contraction rate does not bring the fast variables within a neighbourhood of ϵ/λ of the slow manifold, the fast variables will not have sufficiently relaxed and their dynamics remains stiff, possibly causing numerical instability over the subsequent integration steps.
The bounds for the deviation of the fast variables from the slow manifold are different for PI1 and PI2. In particular, Theorem 5.3 suggests that for a given macrostep size the fast variables deviate less from the slow manifold in our modified version PI2. This will be confirmed numerically in Section 6.

In the next section we prove Theorems 5.2 and 5.3. We formulate the proofs for PI2 and point out where and how they will differ for PI1.

5.1. Error Analysis

We split the error $\mathcal{E}^n$ between the PI approximation of the slow variables and their true value into two parts. Denote by $Y(t^n)$ the time-continuous solution of the reduced ordinary differential equation (2.5) evaluated at time t^n, then

$$\mathcal{E}^n=|y_\varepsilon(t^n)-y^n|$$

$$\leq |y_\varepsilon(t^n) - Y(t^n)| + |y^n - Y(t^n)|,$$

where the first term describes the error between the exact solutions of the full system (2.2)-(2.4) and the reduced slow system (2.5), which we label *reduction error*, with

$$E_r^n = y_\varepsilon(t^n) - Y(t^n), \tag{5.1}$$

and the second term the error between PI and the exact solution of the reduced slow system (2.5), which we label *discretization error*, with

$$E_d^n = y^n - Y(t^n). \tag{5.2}$$

We will bound the two terms separately in the following.

5.2. Reduction error

Setting the initial conditions close to the slow manifold with $y_\varepsilon(0) = Y(0) + c_{0,y}\varepsilon$ and $x_\varepsilon(0) = h_\varepsilon(y_\varepsilon(0)) + c_{0,x}$, we formulate the following theorem for the reduction error E_r^n.

Theorem 5.5. *Given assumptions (A1)–(A3), there exists a constant C_1 such that on a fixed time interval T, for each $t^n \leq T$, the difference between the exact solutions of the reduced and the full system is bounded by*

$$|E_r^n| \leq C_1 \varepsilon,$$

with

$$C_1 = \max(|y_\varepsilon(0) - Y(0)|, L_g |d_\varepsilon(0)|) e^{L_g(1+L_h)t^n},$$

where $d_\varepsilon = x_\varepsilon - h_\varepsilon(y_\varepsilon)$ measures the distance of the fast variables from the slow manifold.

The proof is standard and is omitted here. The interested reader is referred to, for example, [28, 18].

5.3. Discretization error

We bound the discretization error $E_d^n = y^n - Y(t^n)$ in stages. We first give a proof for the convergence of a PI approximation of the reduced dynamics to the true reduced dynamics in Proposition 5.6. We then compare the PI approximations of the reduced and the full multi-scale dynamics, and combine the two results to bound E_d^n.

To achieve the first bound we introduce the auxiliary vector field $\tilde{G}$, which describes the PI2 method applied to the reduced slow system (2.5). We first show that $\tilde{G}(Y(t^n))$ is close to a standard Runge-Kutta solver applied to $Y(t^n + M_1\delta t)$; then we bound the difference between the auxiliary vectorfield $\tilde{G}(y^n)$ and the PI2 vectorfield for the slow variable $\tilde{g}(x^n, y^n)$.

Denote by $\phi^{m,\delta t}$ the flow map for the microsolver of order p applied to the reduced system (2.5) for m microsteps with time step δt. Given initial condition

Y^n at $t = t^n$, we construct $\tilde{G}$ analogously to the construction of $\tilde{g}$ used in PI2. We define $Y_m^{n,j}$ for $j = 1, 2, \ldots, P+1$, $m = 1, 2, \ldots, M_j$ as the output of the microsolver

$$Y_m^{n,j} = \phi^{m,\delta t}\left(Y_0^{n,j}\right), \tag{5.3}$$

analogous to (4.9), with initial condition

$$Y_0^{n,j} = \begin{cases} Y^n & \text{for } j = 1 \\ Y_{M_1}^{n,1} + a_j \Delta t \, G(Y_{M_{j-1}}^{n,j-1}) & \text{for } j > 1 \end{cases}, \tag{5.4}$$

analogous to (4.10) and (4.8). The increments are constructed by

$$K_j(Y^n) = \frac{1}{a_{j+1}}\left(Y_{M_{j+1}}^{n,j+1} - Y_{M_1}^{n,1}\right), \tag{5.5}$$

analogous to (4.11). Combining (5.3)–(5.5), we form the auxiliary vectorfield

$$\tilde{G}(Y^n) = Y_{M_1}^{n,1} - Y^n + \sum_{j=1}^{P} b_j K_j(Y^n), \tag{5.6}$$

analogous to the PI vectorfield (4.19) of the macrosolver.

In the following Proposition we demonstrate that $\tilde{G}$ evaluated at $Y(t^n)$ incurs an error of order $\mathcal{O}(t_\Delta^{P+1})$ over one macrostep, like standard Runge-Kutta methods, with an additional error term incurred by the applications of the microsolver.

Proposition 5.6. *Given assumptions (A1), (A2), (A4) and (A5), $\tilde{G}(Y(t^n))$ provides a numerical estimate of the reduced slow vectorfield with*

$$Y(t^{n+1}) = Y(t^n) + \tilde{G}(Y(t^n)) + \mathcal{O}(t_\Delta^{P+1}, t_\Delta M \delta t),$$

where the error term $\mathcal{O}(t_\Delta^{P+1}, t_\Delta M \delta t)$ is bounded by $C_P^* t_\Delta^{P+1} + C_2^* t_\Delta M \delta t$.

PROOF. The increments $\bar{K}_j$ of a Runge-Kutta solver of order P applied to the reduced system (2.5), initialised at Y with time step t_Δ, are given by

$$\bar{K}_j(Y) = t_\Delta \, G(Y + a_j \bar{K}_{j-1}(Y)). \tag{5.7}$$

For a Runge-Kutta solver of order P initialised at $Y(t^n + M_1 \delta t)$ we have

$$Y(t^{n+1}) = Y(t^n + M_1 \delta t) + \sum_{j=1}^{P} b_j \bar{K}_j(Y(t^n + M_1 \delta t)) + \mathcal{O}\left(t_\Delta^{P+1}\right), \tag{5.8}$$

where the $\mathcal{O}(t_\Delta^{P+1})$ term is bounded by $C_P^* t_\Delta^{P+1}$ [27]. Similarly, a microsolver ϕ of order p satisfies

$$\left|Y(t^n + M_1 \delta t) - Y_{M_1}^{n,1}\right| \leq C_p^* M_1 \delta t^{p+1}. \tag{5.9}$$

The Runge-Kutta solver (5.8) is rewritten as

$$Y(t^{n+1}) = Y_{M_1}^{n,1} + \sum_{j=1}^{P} b_j \bar{K}_j(Y_{M_1}^{n,1}) \tag{5.10}$$

$$+ \sum_{j=1}^{P} b_j \left(\bar{K}_j(Y(t^n + M_1 \delta t)) - \bar{K}_j(Y_{M_1}^{n,1}) \right) + \mathcal{O}\left(t_\Delta^{P+1}, M_1 \delta t^{p+1} \right) .$$

Employing assumptions $(A1)$–$(A2)$ on the Lipschitz continuity of the reduced dynamics, (5.9) and the definition (5.7) of the Runge-Kutta increments $\bar{K}$ we bound

$$\left| \bar{K}_j(Y(t^n + M_1 \delta t)) - \bar{K}_j(Y_{M_1}^{n,1}) \right| \leq L_G t_\Delta \left| Y(t^n + M_1 \delta t) - Y_{M_1}^{n,1} \right|$$

$$+ L_G t_\Delta a_j \left| \bar{K}_{j-1}(Y(t^n + M_1 \delta t)) - \bar{K}_{j-1}(Y_{M_1}^{n,1}) \right|$$

$$\leq L_G t_\Delta C_p^* M_1 \delta t^{p+1}$$

$$+ L_G t_\Delta a_j \left| \bar{K}_{j-1}(Y(t^n + M_1 \delta t)) - \bar{K}_{j-1}(Y_{M_1}^{n,1}) \right| .$$

Iterating this relationship with $a_1 = 0$ yields

$$\left| \bar{K}_j(Y(t^n + M_1 \delta t)) - \bar{K}_j(Y_{M_1}^{n,1}) \right| \leq L_G t_\Delta C_p^* M_1 \delta t^{p+1} + \mathcal{O}(t_\Delta^2 M_1 \delta t^{p+1}) .$$

Upon substitution into (5.10) we obtain

$$Y(t^{n+1}) = Y_{M_1}^{n,1} + \sum_{j=1}^{P} b_j \bar{K}_j(Y_{M_1}^{n,1}) + \mathcal{O}\left(t_\Delta^{P+1}, M_1 \delta t^{p+1} \right) , \tag{5.11}$$

which describes a Runge-Kutta method of order P, initialised at $Y_{M_1}^{n,1}$. The auxiliary vectorfield $\tilde{G}$ given by (5.6) with $Y^n = Y(t^n)$ is now constructed from (5.11). We write

$$Y(t^{n+1}) = Y_{M_1}^{n,1} + \sum_{j=1}^{P} b_j \left(K_j(Y(t^n)) + \bar{K}_j(Y_{M_1}^{n,1}) - K_j(Y(t^n)) \right) + \mathcal{O}\left(t_\Delta^{P+1}, M_1 \delta t^{p+1} \right)$$

$$= Y(t^n) + \tilde{G}(Y(t^n)) + \sum_{j=1}^{P} b_j \left(\bar{K}_j(Y_{M_1}^{n,1}) - K_j(Y(t^n)) \right) + \mathcal{O}\left(t_\Delta^{P+1}, M_1 \delta t^{p+1} \right) , \tag{5.12}$$

where the increments K_j are defined in (5.5).
We now bound $|\bar{K}_j(Y_{M_1}^{n,1}) - K_j(Y(t^n))|$ in (5.12). Rearranging the definition of K_j, (5.5), we obtain

$$Y_{M_1}^{n,1} + a_{j+1} K_j(Y(t^n)) = Y_{M_{j+1}}^{n,j+1}$$

$$= \phi^{M_{j+1},\delta t}\left(Y_0^{n,j+1}\right) . \tag{5.13}$$

Similarly we use the definition of $\bar{K}_j\binom{n,1}{M_1}$, (5.7), to obtain

$$
\begin{aligned}
Y_{M_1}^{n,1} + a_{j+1}\bar{K}_j(Y_{M_1}^{n,1}) &= Y_{M_1}^{n,1} + a_{j+1}t_\Delta\, G\left(Y_{M_1}^{n,1} + a_j\bar{K}_{j-1}\right) \\
&= Y_{M_1}^{n,1} + a_{j+1}\Delta t\, G(Y_{M_1}^{n,1} + a_j K_{j-1}) \\
&\quad + a_{j+1}\Delta t\left(G\left(Y_{M_1}^{n,1} + a_j\bar{K}_{j-1}\right) - G(Y_{M_1}^{n,1} + a_j K_{j-1})\right) \\
&\quad + a_{j+1}M\delta t\, G\left(Y_{M_1}^{n,1} + a_j\bar{K}_{j-1}\right) \\
&= Y_0^{n,j+1} + a_{j+1}M\delta t G(Y_0^{n,j+1}) \tag{5.14} \\
&\quad + a_{j+1}\Delta t\left(G\left(Y_{M_1}^{n,1} + a_j\bar{K}_{j-1}\right) - G(Y_{M_1}^{n,1} + a_j K_{j-1})\right) \\
&\quad + a_{j+1}M\delta t\left(G\left(Y_{M_1}^{n,1} + a_j\bar{K}_{j-1}\right) - G(Y_0^{n,j+1})\right) ,
\end{aligned}
$$

where we have suppressed the dependencies of $\bar{K}_j$ and K_j on the right-hand side and used $t_\Delta = \Delta t + M\delta t$. Subtracting (5.13) from (5.14), applying absolute values and dividing by a_{j+1} yields the bound

$$
\left|\bar{K}_j(Y_{M_1}^{n,1}) - K_j(Y(t^n))\right| \leq \frac{1}{a_{j+1}}\left|Y_0^{n,j+1} + a_{j+1}M\delta t G(Y_0^{n,j+1}) - \phi^{M_{j+1},\delta t}\left(Y_0^{n,j+1}\right)\right|
$$
$$\tag{5.15}$$
$$
+ \Delta t\left|G\left(Y_{M_1}^{n,1} + a_j\bar{K}_{j-1}\right) - G(Y_{M_1}^{n,1} + a_j K_{j-1})\right|
$$
$$
+ M\delta t\left|G\left(Y_{M_1}^{n,1} + a_j\bar{K}_{j-1}\right) - G(Y_0^{n,j+1})\right| .
$$

We now bound the three lines of (5.15) separately. In the first line, we interpret $Y_0^{n,j+1} + a_{j+1}M\delta t G(Y_0^{n,j+1})$ as a single Euler step with time step $a_{j+1}M\delta t$ initialised at $Y_0^{n,j+1}$. The remaining term of the first line, $\phi^{M_{j+1},\delta t}\left(Y_0^{n,j+1}\right)$, describes a microsolver of order p run for $M_{j+1} = a_{j+1}M$ steps, also initialised at $Y_0^{n,j+1}$. Therefore

$$
\left|Y_0^{n,j+1} + a_{j+1}M\delta t G(Y_0^{n,j+1}) - \phi^{M_{j+1},\delta t}\left(Y_0^{n,j+1}\right)\right| \leq \mathcal{O}(a_{j+1}M\delta t^{p+1}, (a_{j+1}M\delta t)^2) .
$$
$$\tag{5.16}$$

The second line in (5.15), employing Assumptions $(A1)$–$(A2)$ on the Lipschitz continuity of the reduced dynamics, is bounded by

$$
\Delta t\left|G\left(Y_{M_1}^{n,1} + a_j\bar{K}_{j-1}\right) - G(Y_{M_1}^{n,1} + a_j K_{j-1})\right| \leq a_j\Delta t L_G\left|\bar{K}_{j-1} - K_{j-1}\right| . \tag{5.17}
$$

Upon using Assumption $(A4)$ on the boundedness of the reduced dynamics and Assumptions $(A1)$–$(A2)$, we bound the term in absolute values in the third line of (5.15) by

$$
\left|G\left(Y_{M_1}^{n,1} + a_j\bar{K}_{j-1}\right) - G(Y_0^{n,j+1})\right| \leq L_G\left|Y_{M_1}^{n,1} + a_j\bar{K}_{j-1} - Y_0^{n,j+1}\right|
$$

$$=L_G \left| a_j t_\Delta G\left(Y_{M_1}^{n,1} + a_{j-1}\bar{K}_{j-2}\right)\right.$$
$$\left. - a_{j+1}\Delta t G\left(Y_{M_j}^{n,j}\right)\right|$$
$$\leq 2C_2^* t_\Delta , \tag{5.18}$$

where we employed the bound on the nodes $a_j \leq 1$, and Assumption $(A5)$ on the smoothness of the reduced dynamics with $C_2^* = \sup|\ddot{Y}| = \sup|DG(Y)G(Y)| = L_G C_G$.

Substituting (5.16)–(5.18) into (5.15) yields the bound

$$\left|\bar{K}_j(Y_{M_1}^{n,1}) - K_j(Y(t^n))\right| \leq a_j \Delta t L_G \left|\bar{K}_{j-1} - K_{j-1}\right| + 2C_2^* t_\Delta M\delta t \tag{5.19}$$
$$+ \mathcal{O}(M\delta t^{p+1}, a_{j+1}(M\delta t)^2)) .$$

Substituting $j = 1$, noting that $a_1 = 0$ and neglecting higher order terms, we have

$$\left|\bar{K}_1(Y_{M_1}^{n,1}) - K_1(Y(t^n))\right| \leq 2C_2^* t_\Delta M\delta t . \tag{5.20}$$

Iteration of (5.19), seeded with (5.20) at $j = 1$, yields

$$\left|\bar{K}_j(Y_{M_1}^{n,1}) - K_j(Y(t^n))\right| \leq 2C_2^* t_\Delta M\delta t .$$

The Proposition now follows directly by substituting into (5.12) and using the weighting condition $\sum_{j=1}^{P} b_j = 1$.

Remark 5.7. *Proposition 5.6 can be readily extended for PI1. This Proposition employs auxiliary increments K_j designed to resemble the PI2 increments (4.11). In order to prove this result for the PI1 method, one should instead employ auxiliary increments $\hat{K}_j = \Delta t G(Y_{M_j}^{n,j})$, which resemble the PI1 increments (4.1). Following from (5.12), one can then readily bound $|\hat{K}_j - \bar{K}_j| \leq 2C_2^* \Delta t M\delta t$ to obtain the same bound.*

Proposition 5.6 establishes that using $\tilde{G}$ to propagate the reduced dynamics incurs error proportional to $t_\Delta^{P+1} + t_\Delta M\delta t$. In particular, these terms do not depend on the order p of the microsolver. To simplify the calculations, we therefore use a forward Euler method as the microsolver for the reduced system. In particular, we consider

$$Y_m^{n,j} = Y_{m-1}^{n,j} + \delta t G(Y_{m-1}^{n,j}) , \tag{5.21}$$

for $0 \leq m \leq M_j$.

We now use Proposition 5.6 to bound the error between the PI2 approximation of the slow variable y^n in a full multiscale simulation and the reduced dynamics $Y(t^n)$.

Lemma 5.8. *Given assumptions (A1)–(A5), the discretization error* $|E_d^n| = |y^n - Y(t^n)|$ *is bounded by*

$$|E_d^n| \le \frac{e^{L_G t^n}}{L_G} \left\{ C_P^* t_\Delta^P + C_2^* M \delta t + \max_{0 \le i \le n-1} \left| \frac{\tilde{g}(x^i, y^i) - \tilde{G}(x^i, y^i)}{t_\Delta} \right| \right\}.$$

PROOF. This result follows from [9]. Employing Proposition 5.6, we have

$$
\begin{aligned}
E_d^n &= E_d^{n-1} + \tilde{g}(x^{n-1}, y^{n-1}) - \tilde{G}(Y(t^{n-1})) + \mathcal{O}(t_\Delta^{P+1}, t_\Delta M \delta t) \\
&= E_d^{n-1} + \mathcal{L}_G^{n-1} E_d^{n-1} + \tilde{g}(x^{n-1}, y^{n-1}) - \tilde{G}(y^{n-1}) + \mathcal{O}(t_\Delta^{P+1}, t_\Delta M \delta t), \quad (5.22)
\end{aligned}
$$

where we used the mean value theorem for vector-valued functions to introduce

$$\mathcal{L}_G^n = \int_0^1 \mathrm{D}\tilde{G}\big(Y(t^n) + \theta(y^n - Y(t^n))\big) \, d\theta, \quad (5.23)$$

where $\mathrm{D}\tilde{G}$ is the Jacobian matrix of $\tilde{G}$. Recall that the $\mathcal{O}(t_\Delta^{P+1}, t_\Delta M \delta t)$ term in (5.22) is bounded by $C_P^* t_\Delta^{P+1} + C_2^* t_\Delta M \delta t$; taking absolute values of (5.22) then yields

$$|E_d^n| \le \big(1 + |\mathcal{L}_G^{n-1}|\big) |E_d^{n-1}| + \big|\tilde{g}(x^{n-1}, y^{n-1}) - \tilde{G}(y^{n-1})\big| + C_P^* t_\Delta^{P+1} + C_2^* t_\Delta M \delta t. \quad (5.24)$$

To bound $|\mathcal{L}_G^{n-1}|$, we first obtain an explicit formula for $\tilde{G}$. Substituting (5.6), (5.5) and the Euler microsolver (5.21) into $\tilde{G}$, (5.6), we obtain

$$
\begin{aligned}
\tilde{G}(Y^n) &= Y_{M_1}^{n,1} - Y^n + \sum_{j=1}^{P} \frac{b_j}{a_{j+1}} \left(Y_{M_{j+1}}^{n,j+1} - Y_{M_1}^{n,1} \right) \\
&= Y^n + \delta t \sum_{k=0}^{M_1-1} G(Y_k^{n,1}) - Y^n \\
&\quad + \sum_{j=1}^{P} \frac{b_j}{a_{j+1}} \left(Y_0^{n,j+1} + \delta t \sum_{k=0}^{M_{j+1}-1} G(Y_k^{n,j+1}) - Y_{M_1}^{n,1} \right).
\end{aligned}
$$

Substituting $Y_0^{n,j+1} = Y_{M_1}^{n,1} + a_{j+1} \Delta t \, G(Y_{M_j}^{n,j})$ from (5.4), we obtain the explicit formula

$$\tilde{G}(Y^n) = \delta t \sum_{k=0}^{M_1-1} G(Y_k^{n,1}) + \sum_{j=1}^{P} \frac{b_j}{a_{j+1}} \left(a_{j+1} \Delta t \, G(Y_{M_j}^{n,j}) + \delta t \sum_{k=0}^{M_{j+1}-1} G(Y_k^{n,j+1}) \right).$$

Substituting into (5.23) with $Y^n = Y(t^n) + \theta(y^n - Y(t^n))$, taking absolute values and employing Assumptions (A1)–(A2) on the Lipshitz constant L_G of the reduced dynamics yields

$$|\mathcal{L}_G^n| \le \int_0^1 \left| \mathrm{D}\tilde{G}\big(Y(t^n) + \theta(y^n - Y(t^n))\big) \right| d\theta$$

$$\leq \int_0^1 \delta t \sum_{k=0}^{M_1-1} \left| \mathrm{D}G(Y_k^{n,1}) \right| + \sum_{j=1}^{P} \frac{b_j}{a_{j+1}} a_{j+1} \Delta t \left| \mathrm{D}G(Y_{M_j}^{n,j}) \right|$$

$$+ \delta t \sum_{k=0}^{M_{j+1}-1} \left| \mathrm{D}G(Y_k^{n,j+1}) \right| d\theta$$

$$\leq \int_0^1 \left(L_G M_1 \delta t + \sum_{j=1}^{P} \frac{b_j}{a_{j+1}} \left(L_G a_{j+1} \Delta t + L_G M_{j+1} \delta t \right) \right) d\theta \ .$$

Recalling $M_j = a_j M$ for $j > 1$, and using the weighting condition $\sum_{j=1}^{P} b_j = 1$, we obtain the bound

$$|\mathcal{L}_G^n| \leq L_G \left(t_\Delta + M_1 \delta t \right) \ .$$

We substitute this bound into (5.24), with

$$|E_d^n| \leq (1 + L_G(t_\Delta + M_1 \delta t)) |E_d^{n-1}| + \left| \tilde{g}(x^{n-1}, y^{n-1}) - \tilde{G}(y^{n-1}) \right|$$

$$+ C_P^* t_\Delta^{P+1} + C_2^* t_\Delta M \delta t \ .$$

Iterating the recursive relationship with $E_d^0 = 0$ yields

$$|E_d^n| \leq \sum_{m=0}^{n-1} (1 + L_G(t_\Delta + M_1 \delta t))^m \left(C_P^* t_\Delta^{P+1} + C_2^* t_\Delta M \delta t \right.$$

$$\left. + \max_{0 \leq i \leq n-1} \frac{|\tilde{g}(x^i, y^i) - \tilde{G}(y^i)|}{t_\Delta} \right)$$

$$\leq \frac{e^{L_G t^n}}{L_G} \left(C_P^* t_\Delta^P + C_2^* M \delta t + \max_{0 \leq i \leq n-1} \frac{|\tilde{g}(x^i, y^i) - \tilde{G}(y^i)|}{t_\Delta} \right) \ ,$$

where $t^n = n(t_\Delta + M_1 \delta t)$.

Lemma 5.8 establishes that the error between the PI2 approximation of the slow variables of the full multiscale system (2.1) and the true reduced dynamics (2.5) contains a term proportional to the order of the macrosolver t_Δ^P, a term proportional to the length of the microsolver $M\delta t$ and an additional term proportional to $\max_{0 \leq i \leq n-1} |\tilde{g}(x^{n-1}, y^{n-1}) - \tilde{G}(y^{n-1})|$. The latter term measures the difference between the PI2 vector field $\tilde{g}$ of the slow variables, and the auxiliary vector field $\tilde{G}$ initialised at the same point y^{n-1}. In order to bound this term we define the deviation of the PI approximation of the fast variables from the approximate slow manifold over the increments,

$$d_m^{n,j} = x_m^{n,j} - h_0(y_m^{n,j}), \tag{5.25}$$

for $j = 1, 2, \ldots, P+1$, $m = 1, 2, \ldots, M_j$. The following Lemma bounds $|d_m^{n,j}|$.

Lemma 5.9. *Given Assumptions (A1) and (A4), the error between the fast variables and the approximate slow manifold during the application of a micro-solver of order p is bounded for all $0 \le m \le M_j$ by*

$$|d_m^{n,j}| \le \rho^m \left(-\frac{\delta t}{\varepsilon}\right)|d_0^{n,j}| + L_h C_g \varepsilon \ ,$$

where

$$\rho\left(-\frac{\delta t}{\varepsilon}\right) = \sum_{k=0}^{p} \frac{\left(-\frac{\delta t}{\varepsilon}\right)^k}{k!} \ .$$

The first term in the Lemma is a manifestation of the attraction of the fast variables towards the slow manifold along their stable eigendirection. The second term proportional to ε describes, as we will see below, the cumulative drift of the slow variables y during the microsteps causing a departure from the slow manifold for nonconstant $h_0(y)$.

To ensure convergence of the fast variables to the approximate slow manifold we require

$$0 \le \rho\left(-\frac{\delta t}{\varepsilon}\right) < 1 \ .$$

PROOF. Denote the increments of the microsolver as $\bar{k}_{x,i}$ and $\bar{k}_{y,i}$ for the fast and slow variables respectively, with nodes a_i' and weights b_i'. Employing the fast vector field (2.4) we write the fast increments $\bar{k}_{x,i}$ analogously to (3.1) as

$$\bar{k}_{x,i}(x_m^{n,j}, y_m^{n,j}) = -\frac{\Lambda \delta t}{\varepsilon}\left(x_m^{n,j} + a_i'\bar{k}_{x,i-1}\right) + \frac{\Lambda \delta t}{\varepsilon}h_0(y_m^{n,j} + a_i'\bar{k}_{y,i-1})$$

$$= -\frac{\Lambda \delta t}{\varepsilon}\left(x_m^{n,j} - h_0(y_m^{n,j}) + a_i'\bar{k}_{x,i-1}\right) + \mathcal{O}\left(\frac{\delta t^2}{\varepsilon}\right)$$

$$= -\frac{\Lambda \delta t}{\varepsilon}\left(d_m^{n,j} + a_i'\bar{k}_{x,i-1}\right) + \mathcal{O}\left(\frac{\delta t^2}{\varepsilon}\right) \ , \qquad (5.26)$$

where we have used that $\bar{k}_{y,i} = \mathcal{O}(\delta t)$. Introducing the increment associated with a linear system $\bar{k}_{\mathrm{lin},i}(d_m^{n,j}) = -\Lambda \delta t\left(d_m^{n,j} + a_i'\bar{k}_{\mathrm{lin},i-1}\right)/\varepsilon$, we write

$$\bar{k}_{x,i}(x_m^{n,j}, y_m^{n,j}) - \bar{k}_{\mathrm{lin},i}(d_m^{n,j}) = -\frac{\Lambda \delta t}{\varepsilon}a_i'\left(\bar{k}_{x,i-1} - \bar{k}_{\mathrm{lin},i-1}\right) + \mathcal{O}\left(\frac{\delta t^2}{\varepsilon}\right)$$

$$= \mathcal{O}\left(\frac{\delta t^2}{\varepsilon}\right) \ .$$

The linear component $\bar{k}_{\mathrm{lin},i}$ of the increment can be interpreted as the increment of the microsolver applied to the linear system $\dot{d}_\varepsilon = -\Lambda d_\varepsilon/\varepsilon$ with initial condition $d_m^{n,j}$. Therefore, as discussed in Section 3, a microstep taken with the linear increments $\bar{k}_{\mathrm{lin},i}$ can be written as

$$d_m^{n,j} + \sum_{i=1}^{p} b_i'\bar{k}_{\mathrm{lin},i}(d_m^{n,j}) = \rho\left(-\frac{\Lambda \delta t}{\varepsilon}\right)d_m^{n,j} \ .$$

Employing (3.2) the microstep is expressed as

$$x_{m+1}^{n,j} = x_m^{n,j} + \sum_{i=1}^{p} b_i' \bar{k}_{x,i}(x_m^{n,j}, y_m^{n,j})$$

$$= x_m^{n,j} + \sum_{i=1}^{p} b_i' \bar{k}_{\text{lin},i}(d_m^{n,j}) + \mathcal{O}\left(\frac{\delta t^2}{\varepsilon}\right)$$

$$= d_m^{n,j} + \sum_{i=1}^{p} b_i' \bar{k}_{\text{lin},i}(d_m^{n,j}) + h_0(y_m^{n,j}) + \mathcal{O}\left(\frac{\delta t^2}{\varepsilon}\right)$$

$$= \rho\left(-\frac{\Lambda \delta t}{\varepsilon}\right) d_m^{n,j} + h_0(y_m^{n,j}) + \mathcal{O}\left(\frac{\delta t^2}{\varepsilon}\right).$$

Then

$$\left| d_{m+1}^{n,j} \right| = \left| x_{m+1}^{n,j} - h_0(y_{m+1}^{n,j}) \right|$$

$$\leq \left| \rho\left(-\frac{\Lambda \delta t}{\varepsilon}\right) \right| \left| d_m^{n,j} \right| + \left| h_0(y_m^{n,j}) - h_0(y_{m+1}^{n,j}) \right| + \mathcal{O}\left(\frac{\delta t^2}{\varepsilon}\right)$$

$$\leq \left| \rho\left(-\frac{\Lambda \delta t}{\varepsilon}\right) \right| \left| d_m^{n,j} \right| + L_h C_g \delta t + \mathcal{O}\left(\frac{\delta t^2}{\varepsilon}\right). \qquad (5.27)$$

The first term in this bound represents the rate of convergence of the fast variables to the approximate slow manifold; for stability we require $\left|\rho\left(-\frac{\Lambda \delta t}{\varepsilon}\right)\right| < 1$. The term $L_h C_g \delta t$ stems from the drift in the slow variables over a microstep. The slowest rate of convergence to the slow manifold is given by $\min(\lambda_{ii}) = 1$, so we obtain

$$\left| \rho\left(-\frac{\Lambda \delta t}{\varepsilon}\right) \right| \leq \rho\left(-\frac{\delta t}{\varepsilon}\right)$$

$$= 1 - \frac{\delta t}{\varepsilon} + \frac{1}{2}\left(-\frac{\delta t}{\varepsilon}\right)^2 + \cdots + \frac{1}{p!}\left(-\frac{\delta t}{\varepsilon}\right)^p$$

$$< 1.$$

Iterating (5.27) then yields

$$\left| d_m^{n,j} \right| \leq \rho^m\left(-\frac{\delta t}{\varepsilon}\right) \left| d_m^{n,j} \right| + \frac{L_h C_g \delta t + \mathcal{O}\left(\frac{\delta t^2}{\varepsilon}\right)}{1 - \left|\rho\left(-\frac{\delta t}{\varepsilon}\right)\right|}$$

$$\leq \rho^m\left(-\frac{\delta t}{\varepsilon}\right) \left| d_m^{n,j} \right| + L_h C_g \varepsilon + \mathcal{O}(\delta t),$$

completing the proof of the Lemma.

Remark 5.10. *In the above Lemma we Taylor expand the terms in $h_0(y_m^{n,j} + a_i' \bar{k}_{y,i-1})$ up to $\mathcal{O}(\delta t^2/\varepsilon)$. However, higher-order terms may improve the error*

bound. For instance, one can show that for a fourth-order Runge-Kutta micro-solver,

$$\left|d_{m+1}^{n,j}\right| \leq \rho\left(-\frac{\delta t}{\varepsilon}\right)\left|d_m^{n,j}\right| + \frac{9}{24}L_h C_g \delta t \ .$$

Remark 5.11. *Optimal convergence of the fast variables to the approximate slow manifold during the application of the microsolver is given by a forward Euler microsolver with $0 < \delta t \leq \frac{2\varepsilon}{\lambda+1}$, where*

$$\rho\left(-\frac{\delta t}{\varepsilon}\right) = 1 - \frac{\delta t}{\varepsilon} \ ,$$

and the convergence rate $\rho^m(-\delta t/\varepsilon)$ is bounded above by the exponential convergence $\exp(-m\delta t/\varepsilon)$. The full stability region for the Euler microsolver is $0 < \delta t < 2\varepsilon/\lambda$; for further details, see [18].

From Lemma 5.9 it follows that the rate of convergence of the fast variables to the approximate slow manifold is optimal for an Euler microsolver, and in Lemma 5.8 we demonstrated that the dominant error terms between the PI2 approximation y^n and the true reduced dynamics $Y(t^n)$ do not depend on the order p of the microsolver. We therefore choose as the microsolver for the PI2 scheme the forward Euler method to simplify the calculations, and write

$$y_m^{n,j} = y_{m-1}^{n,j} + \delta t\, g(x_{m-1}^{n,j}, y_{m-1}^{n,j}) \ , \tag{5.28}$$

$$x_m^{n,j} = x_{m-1}^{n,j} + \delta t\, f(x_{m-1}^{n,j}, y_{m-1}^{n,j}) \ , \tag{5.29}$$

We now bound the distance of the PI approximation $y_m^{n,j}$ of the slow variables over the microsteps of the full system (2.1), from $Y_m^{n,j}$, the PI approximation of the reduced dynamics over the microsteps.

Lemma 5.12. *Assuming (A1),(A2) and (A6), the PI2 numerical estimate $y_{M_j}^{n,j}$ of the slow variable after the application of the microsolver at the j-th increment is close to the numerical estimate $Y_{M_j}^{n,j}$ of the reduced slow variable which was initialized at $Y_0^{n,1} = y^n$, with*

$$|y_{M_j}^{n,j} - Y_{M_j}^{n,j}| \leq 2L_g\left(3\varepsilon + a_j\Delta t\left(1 - \frac{\delta t}{\varepsilon}\right)^{aM}\right)\max_{1\leq k\leq j}\left|d_0^{n,k}\right| + 2L_g L_h C_g a_j t_\Delta \varepsilon$$

$$+ \mathcal{O}\left(\left(\Delta t\varepsilon + \Delta t^2(1 - \frac{\delta t}{\varepsilon})^{aM}\right)\max_{1\leq k\leq j}\left|d_0^{n,k}\right|, t_\Delta^2\varepsilon, M\delta t\varepsilon, M_1\delta t\varepsilon\right) \ ,$$

for $2 \leq j \leq P$, where $aM = \min_{j>1} M_j$, and

$$|y_{M_1}^{n,1} - Y_{M_1}^{n,1}| \leq 2L_g\varepsilon\left|d_0^{n,1}\right| + \mathcal{O}(M_1\delta t\varepsilon) \ ,$$

for $j = 1$.

PROOF. Employing the definition of the Euler microsolvers (5.28) for the PI2 scheme and (5.21) for the reduced scheme, and Assumptions $(A1)$–$(A2)$ on the Lipschitz continuity of the reduced dynamics gives

$$
\begin{aligned}
\left| y_{M_j}^{n,j} - Y_{M_j}^{n,j} \right| &\leq \left| y_{M_j-1}^{n,j} - Y_{M_j-1}^{n,j} \right| + \delta t \left| g(x_{M_j-1}^{n,j}, y_{M_j-1}^{n,j}) - G(Y_{M_j-1}^{n,j}) \right| \\
&\leq \left| y_{M_j-1}^{n,j} - Y_{M_j-1}^{n,j} \right| + \delta t \left| g(x_{M_j-1}^{n,j}, y_{M_j-1}^{n,j}) - G(y_{M_j-1}^{n,j}) \right| \\
&\quad + \delta t \left| G(y_{M_j-1}^{n,j}) - G(Y_{M_j-1}^{n,j}) \right| \\
&\leq (1 + L_G \delta t) \left| y_{M_j-1}^{n,j} - Y_{M_j-1}^{n,j} \right| \\
&\quad + \delta t \left| g(x_{M_j-1}^{n,j}, y_{M_j-1}^{n,j}) - g(h_\varepsilon(y_{M_j-1}^{n,j}), y_{M_j-1}^{n,j}) \right| \\
&\leq (1 + L_G \delta t) \left| y_{M_j-1}^{n,j} - Y_{M_j-1}^{n,j} \right| + L_g \delta t \left| x_{M_j-1}^{n,j} - h_\varepsilon(y_{M_j-1}^{n,j}) \right| \\
&= (1 + L_G \delta t) \left| y_{M_j-1}^{n,j} - Y_{M_j-1}^{n,j} \right| + L_g \delta t \left| d_{M_j-1}^{n,j} \right| \\
&\quad + L_g \delta t \left| h_0(y_{M_j-1}^{n,j}) - h_\varepsilon(y_{M_j-1}^{n,j}) \right| \\
&= (1 + L_G \delta t) \left| y_{M_j-1}^{n,j} - Y_{M_j-1}^{n,j} \right| + L_g \delta t \left| d_{M_j-1}^{n,j} \right| + L_g L_\varepsilon \varepsilon \delta t + \mathcal{O}(\delta t \varepsilon^2) \,,
\end{aligned}
$$

where we have defined $h_0(y) - h_\varepsilon(y) = L_\varepsilon \varepsilon + \mathcal{O}(\varepsilon^2)$. Employing Lemma 5.9 on $|d_{M_j-1}^{n,j}|$ for a forward Euler microsolver yields the recursive bound

$$
\begin{aligned}
\left| y_{M_j}^{n,j} - Y_{M_j}^{n,j} \right| &\leq (1 + L_G \delta t) \left| y_{M_j-1}^{n,j} - Y_{M_j-1}^{n,j} \right| \\
&\quad + L_g \delta t \left(1 - \frac{\delta t}{\varepsilon} \right)^{M_j-1} \left| d_0^{n,j} \right| + L_g (L_h C_g + L_\varepsilon) \varepsilon \delta t \,,
\end{aligned}
$$

which, upon iterating, gives

$$
\begin{aligned}
\left| y_{M_j}^{n,j} - Y_{M_j}^{n,j} \right| &\leq (1 + L_G \delta t)^{M_j} \left| y_0^{n,j} - Y_0^{n,j} \right| \\
&\quad + L_g \delta t \left| d_0^{n,j} \right| \sum_{k=0}^{M_j-1} \left(1 - \frac{\delta t}{\varepsilon} \right)^k (1 + L_G \delta t)^{M_j-1-k} \\
&\quad + L_g (L_h C_g + L_\varepsilon) \varepsilon \delta t \sum_{k=0}^{M_j-1} (1 + L_G \delta t)^k \\
&= (1 + L_G \delta t)^{M_j} \left| y_0^{n,j} - Y_0^{n,j} \right| \\
&\quad + L_g \delta t \left| d_0^{n,j} \right| \frac{(1 + L_G \delta t)^{M_j} - \left(1 - \frac{\delta t}{\varepsilon} \right)^{M_j}}{L_G \delta t + \frac{\delta t}{\varepsilon}} \\
&\quad + L_g (L_h C_g + L_\varepsilon) \varepsilon \delta t \frac{(1 + L_G \delta t)^{M_j} - 1}{L_G \delta t} \\
&\leq e^{L_G M_j \delta t} \left| y_0^{n,j} - Y_0^{n,j} \right| + L_g \varepsilon \left| d_0^{n,j} \right| \frac{e^{L_G M_j \delta t}}{L_G \varepsilon + 1}
\end{aligned}
$$

$$+ L_g(L_h C_g + L_\varepsilon)\varepsilon\,\delta t\,\frac{e^{L_G M_j \delta t} - 1}{L_G \delta t}\;.$$

Realising that $e^{L_G M_j \delta t} - 1 \le 2 L_G M_j \delta t$ under Assumption ($A6$), we obtain

$$\left| y_{M_j}^{n,j} - Y_{M_j}^{n,j} \right| \le 2 \left| y_0^{n,j} - Y_0^{n,j} \right| + 2 L_g \varepsilon \left| d_0^{n,j} \right| + \mathcal{O}(\varepsilon M_j \delta t)\;. \tag{5.30}$$

At $j = 1$ we initialize at $y_0^{n,1} = Y_0^{n,1} = y^n$, obtaining the desired bound

$$\left| y_{M_1}^{n,1} - Y_{M_1}^{n,1} \right| \le 2 L_g \varepsilon \left| d_0^{n,1} \right| + \mathcal{O}(\varepsilon M_1 \delta t)\;. \tag{5.31}$$

For $j > 1$, we have

$$\begin{aligned}
y_0^{n,j} &= y_{M_1}^{n,1} + a_j \hat{k}_{y,j-1}\\
&= y_{M_1}^{n,1} + a_j \Delta t\, g(x_{M_{j-1}}^{n,j-1}, y_{M_{j-1}}^{n,j-1})\;,
\end{aligned}$$

using the definitions (4.10) and (4.16), and

$$Y_0^{n,j} = Y_{M_1}^{n,1} + a_j \Delta t\, G(Y_{M_{j-1}}^{n,j-1})\;,$$

using (5.4). Substituting these into (5.30) and employing assumptions *(A1)–(A2)*, we obtain

$$\begin{aligned}
\left| y_{M_j}^{n,j} - Y_{M_j}^{n,j} \right| \le\; &2 \left| y_{M_1}^{n,1} - Y_{M_1}^{n,1} \right| + 2 a_j \Delta t \left| g(x_{M_{j-1}}^{n,j-1}, y_{M_{j-1}}^{n,j-1}) - G(Y_{M_{j-1}}^{n,j-1}) \right|\\
&+ 2 L_g \varepsilon \left| d_0^{n,j} \right| + \mathcal{O}(M \delta t \varepsilon)\\
\le\; &4 L_g \varepsilon \left| d_0^{n,1} \right| + 2 a_j \Delta t \left| g(x_{M_{j-1}}^{n,j-1}, y_{M_{j-1}}^{n,j-1}) - G(y_{M_{j-1}}^{n,j-1}) \right|\\
&+ 2 a_j \Delta t \left| G(y_{M_{j-1}}^{n,j-1}) - G(Y_{M_{j-1}}^{n,j-1}) \right| + 2 L_g \varepsilon \left| d_0^{n,j} \right| + \mathcal{O}(M \delta t \varepsilon, M_1 \delta t \varepsilon)\\
\le\; &4 L_g \varepsilon \left| d_0^{n,1} \right| + 2 L_g a_j \Delta t \left| d_{M_{j-1}}^{n,j-1} \right| + 2 L_G a_j \Delta t \left| y_{M_{j-1}}^{n,j-1} - Y_{M_{j-1}}^{n,j-1} \right|\\
&+ 2 L_g \varepsilon \left| d_0^{n,j} \right| + \mathcal{O}(M \delta t \varepsilon, M_1 \delta t \varepsilon)\\
\le\; &4 L_g \varepsilon \left| d_0^{n,1} \right| + 2 L_g a_j \Delta t \left(1 - \frac{\delta t}{\varepsilon} \right)^{M_j} \left| d_0^{n,j-1} \right| + 2 L_g L_h C_g a_j \Delta t \varepsilon\\
&+ 2 L_G a_j \Delta t \left| y_{M_{j-1}}^{n,j-1} - Y_{M_{j-1}}^{n,j-1} \right| + 2 L_g \varepsilon \left| d_0^{n,j} \right| + \mathcal{O}(M \delta t \varepsilon, M_1 \delta t \varepsilon)\;,
\end{aligned}$$

where we have employed (5.31) to bound $\left| y_{M_1}^{n,1} - Y_{M_1}^{n,1} \right|$ and Lemma 5.9 to bound $|d_{M_{j-1}}^{n,j-1}|$. Rearranging and taking the maximum over all increments in the terms in $|d_0^{n,j}|$ and $|d_0^{n,1}|$ produces

$$\left| y_{M_j}^{n,j} - Y_{M_j}^{n,j} \right| \le 2 L_g \left(3\varepsilon + a_j \Delta t \left(1 - \frac{\delta t}{\varepsilon} \right)^{aM} \right) \max_{1 \le k \le j} \left| d_0^{n,k} \right| + 2 L_g L_h C_g a_j t_\Delta \varepsilon$$

$$+2L_G a_j \Delta t \left| y_{M_{j-1}}^{n,j-1} - Y_{M_{j-1}}^{n,j-1} \right| + \mathcal{O}(M\delta t\varepsilon, M_1 \delta t\varepsilon) \ .$$

Iterating this relation yields to lowest order

$$\left| y_{M_j}^{n,j} - Y_{M_j}^{n,j} \right| \leq 2L_g \left(3\varepsilon + a_j \Delta t \left(1 - \frac{\delta t}{\varepsilon} \right)^{aM} \right) \max_{1 \leq k \leq j} \left| d_0^{n,k} \right| + 2L_g L_h C_g a_j t_\Delta \varepsilon$$
$$+ \mathcal{O}\left(\left(\Delta t\varepsilon + \Delta t^2 (1 - \frac{\delta t}{\varepsilon})^{aM} \right) \max_{1 \leq k \leq j} \left| d_0^{n,k} \right|, t_\Delta^2 \varepsilon, M\delta t\varepsilon, M_1 \delta t\varepsilon \right) \ .$$

Lemma 5.12 provides bounds on the difference between solutions of the PI approximation of the slow variables in the full multiscale system and those of the PI approximation of the reduced system during the application of the microsolver. We use this result to bound the difference between the vectorfield of the PI2 method $\tilde{g}$ given by (4.19) and the auxiliary vectorfield $\tilde{G}$ given by (5.6).

Lemma 5.13. *Assuming (A1)-(A6), the auxiliary vectorfield $\tilde{G}$ is close to the vectorfield $\tilde{g}$ with*

$$|\tilde{g}(x^n, y^n) - \tilde{G}(y^n)| \leq 2L_g \left(\frac{5}{a}\varepsilon + \Delta t \left(1 - \frac{\delta t}{\varepsilon} \right)^{aM} \right) \max_{1 \leq k \leq P+1} |d_0^{n,k}| + 2L_g L_h C_g t_\Delta \varepsilon$$
$$+ \mathcal{O}(M_1 \delta t\varepsilon) \ ,$$

PROOF. We write

$$|\tilde{g}(x^n, y^n) - \tilde{G}(y^n)| = \left| y_{M_1}^{n,1} - Y_{M_1}^{n,1} + \sum_{j=1}^{P} \frac{b_j}{a_{j+1}} \left(y_{M_{j+1}}^{n,j+1} - y_{M_1}^{n,1} - (Y_{M_{j+1}}^{n,j+1} - Y_{M_1}^{n,1}) \right) \right|$$
$$\leq \left(\frac{1}{a} - 1 \right) \left| y_{M_1}^{n,1} - Y_{M_1}^{n,1} \right| + \sum_{j=1}^{P} \frac{b_j}{a_{j+1}} \left| y_{M_{j+1}}^{n,j+1} - Y_{M_{j+1}}^{n,j+1} \right| \ ,$$

where we have used that $1 < 1/a_j < 1/a$ for $j > 1$. Employing Lemma 5.12 yields

$$|\tilde{g}(x^n, y^n) - \tilde{G}(y^n)| \leq (\frac{1}{a} - 1) \left(2L_g\varepsilon \left| d_0^{n,1} \right| + 2L_g L_h C_g M_1 \delta t\varepsilon \right)$$
$$+ \sum_{j=1}^{P} b_j \left(2L_g \left(\frac{3}{a_{j+1}}\varepsilon + \Delta t \left(1 - \frac{\delta t}{\varepsilon} \right)^{aM} \right) \max_{1 \leq k \leq j} \left| d_0^{n,k} \right| \right.$$
$$\left. + 2L_g L_h C_g t_\Delta \varepsilon \right)$$
$$\leq 2L_g \left((\frac{4}{a} - 1)\varepsilon + \Delta t \left(1 - \frac{\delta t}{\varepsilon} \right)^{aM} \right) \max_{1 \leq k \leq P+1} |d_0^{n,k}|$$
$$+ 2L_g L_h C_g t_\Delta \varepsilon + 2(1 - \frac{1}{a}) L_g L_h C_g M_1 \delta t\varepsilon \ .$$

Remark 5.14. *In order to prove the above result for the PI1 method, one constructs $\tilde{G}$ with the increments given by vector field evaluations as discussed in Remark 5.7. The Lemma then follows along the same lines as the proofs in [11, 18].*

We are now in the position to establish the bound on the discretization error

$$|E_d^n| = |y^n - Y(t^n)|,$$

which we formulate in the following theorem.

Theorem 5.15. *Given assumptions (A1)–(A6), there exists a constant C such that on a fixed time interval T, for each $n\Delta t \leq T$, the error between the solution of the projective integration scheme PI2 and the exact solutions of the reduced system is bounded by*

$$|E_d^n| \leq C \left(t_\Delta^P + M\delta t + \left(\frac{\varepsilon}{t_\Delta} + e^{-\frac{aM\delta t}{\varepsilon}} \right) |d_{\max}^n| + \varepsilon \right),$$

where $|d_{\max}^n| := \max_{\substack{0 \leq i < n, \\ 1 \leq k \leq P+1}} |d_0^{i,k}|$ is the maximal deviation of the fast variables from the approximate slow manifold over the increments and macrosteps.

PROOF. Combining the bound on the discretization error obtained in Proposition 5.6,

$$|E_d^n| \leq \frac{e^{L_G t^n}}{L_G} \left\{ C_P^* t_\Delta^P + C_2^* M\delta t + \max_{0 \leq i \leq n-1} \left| \frac{\tilde{g}(x^i, y^i) - \tilde{G}(x^i, y^i)}{t_\Delta} \right| \right\},$$

with Lemma 5.13 we obtain

$$|E_d^n| \leq \frac{e^{L_G t^n}}{L_G} \left\{ C_P^* t_\Delta^P + C_2^* M\delta t + 2L_g \left(\frac{5}{a} \frac{\varepsilon}{t_\Delta} + e^{-\frac{aM\delta t}{\varepsilon}} \right) |d_{\max}^n| + 2L_g L_h C_g \varepsilon \right\}.$$

Theorem 5.2 now follows from Theorems 5.5 and 5.15.

Remark 5.16. *Following the comments in Remarks 5.7 and 5.14, one can obtain the same bound for the PI1 method as obtained in Theorem 5.15 for the PI2 method.*

Besides the parameters used in the numerical scheme, i.e. the macrostep size t_Δ, the number of microsteps M with microstep size δt, and the time scale parameter ε, the error bound also involves the maximal deviation of the fast variables from the approximate slow manifold $|d_{\max}^n|$.

We now establish Theorem 5.3 by bounding the distance of the fast variables from the slow manifold over the increments and macrosteps in the PI1 and PI2 schemes. This provides stability conditions for the fast variables in the PI1 and PI2 schemes, i.e. conditions under which $|d_{\max}^n|$ is finite. These stability conditions are crucial for the successful application of the seamless PI methods, since - as we shall see - the fast variables depart from the slow manifold at rate proportional to $\lambda \Delta t / \varepsilon \gg 1$ over the increments and macrosteps.

5.4. Stability of the fast variables

We first bound the distance of the fast variables from the slow manifold over the PI1 increments $\hat{k}_j$, which are employed in both PI1 and PI2.

Lemma 5.17. *Given assumptions (A1) and (A4), the distance of the fast variables from the approximate slow manifold after the j-th increment of the n-th macrostep in the PI1 and PI2 methods, given by $|d_0^{n,j}| = |x_0^{n,j} - h_0(y_0^{n,j})|$, satisfies*

$$|d_0^{n,j}| \leq \left(\frac{\lambda \Delta t}{\varepsilon} \left(1 - \frac{\delta t}{\varepsilon} \right)^{aM} \right)^j |d^n| + L_h C_g (1+\lambda) \Delta t \sum_{k=0}^{j-1} \left(\frac{\lambda \Delta t}{\varepsilon} \left(1 - \frac{\delta t}{\varepsilon} \right)^{aM} \right)^k$$
$$+ \mathcal{O}\left(\varepsilon, \left(1 - \frac{\delta t}{\varepsilon} \right)^{M_1} |d^n| \right) ,$$

where $d_0^{n,1} = d^n = x^n - h_0(y^n)$, and where we define $aM = \min_j M_j$ for the PI1 method to preserve the notation for both methods.

The first term in this result measures the combined effect of the convergence of the fast variables towards the slow manifold over the microsteps, proportional to $(1 - \delta t / \varepsilon)^{aM}$, and the departure of the fast variables from the slow manifold over the subsequent increment, proportional to $\lambda \Delta t / \varepsilon$.

PROOF. The PI1 and PI2 methods (cf (4.3) and (4.10)), respectively satisfy

$$x_0^{n,j+1} = x_{M_1}^{n,1} + a_{j+1} \hat{k}_{x,j}(x^n, y^n)$$
$$= x_{M_1}^{n,1} + a_{j+1} \Delta t \frac{\Lambda}{\varepsilon} (-x_{M_j}^{n,j} + h_0(y_{M_j}^{n,j})) .$$

Then

$$\left| d_0^{n,j+1} \right| = \left| x_0^{n,j+1} - h_0(y_0^{n,j+1}) \right|$$
$$= \left| x_{M_1}^{n,1} + a_{j+1} \Delta t \frac{\Lambda}{\varepsilon} (-x_{M_j}^{n,j} + h_0(y_{M_j}^{n,j})) - h_0(y_0^{n,j+1}) \right|$$
$$= \left| d_{M_1}^{n,1} - a_{j+1} \Delta t \frac{\Lambda}{\varepsilon} d_{M_j}^{n,j} + h_0(y_{M_1}^{n,1}) - h_0(y_0^{n,j+1}) \right|$$
$$\leq \frac{\lambda \Delta t}{\varepsilon} \left| d_{M_j}^{n,j} \right| + L_h \left| y_{M_1}^{n,1} - y_0^{n,j+1} \right| + \left| d_{M_1}^{n,1} \right| . \tag{5.32}$$

Employing Lemma 5.9 with the Euler microsolver and substituting the slow component of (4.3) and the definition of $\hat{k}$, (4.5) and (4.16) respectively, for $y_0^{n,j+1}$ we obtain

$$\left|d_0^{n,j+1}\right| \leq \frac{\lambda\Delta t}{\varepsilon}\left(\left(1-\frac{\delta t}{\varepsilon}\right)^{M_j}\left|d_0^{n,j}\right|+L_hC_g\varepsilon\right)+L_hC_g\Delta t+\left(1-\frac{\delta t}{\varepsilon}\right)^{M_1}|d^n|+L_hC_g\varepsilon$$

$$\leq \frac{\lambda\Delta t}{\varepsilon}\left(1-\frac{\delta t}{\varepsilon}\right)^{aM}\left|d_0^{n,j}\right|+L_hC_g\left(1+\lambda\right)\Delta t+\left(1-\frac{\delta t}{\varepsilon}\right)^{M_1}|d^n|+L_hC_g\varepsilon .$$

Iterating this recursive relation concludes the Lemma.

We now prove Theorem 5.3 in two parts. We first establish bounds on the distance of the fast variables from the slow manifold after one macrostep of the PI1 scheme in Lemma 5.18, and then follow with the corresponding bound for the PI2 scheme in Lemma 5.20.

Lemma 5.18. *Given assumptions (A1) and (A4), the distance of the fast variables from the approximate slow manifold after the n-th macrostep in the PI1 scheme, given by $|d^{n+1}|=|x^{n+1}-h_0(y^{n+1})|$, satisfies the recurrence relation*

$$|d^{n+1}|\leq\sum_{j=1}^{P}b_j\left(\frac{\lambda\Delta t}{\varepsilon}\left(1-\frac{\delta t}{\varepsilon}\right)^{aM}\right)^j|d^n|$$

$$+L_hC_g(1+\lambda)\Delta t\sum_{j=0}^{P-1}\left(\frac{\lambda\Delta t}{\varepsilon}\left(1-\frac{\delta t}{\varepsilon}\right)^{aM}\right)^j+\mathcal{O}\left(\varepsilon,\left(1-\frac{\delta t}{\varepsilon}\right)^{M_1}|d^n|\right) .$$

In particular, the fast variables do not diverge if

$$\frac{\lambda\Delta t}{\varepsilon}\left(1-\frac{\delta t}{\varepsilon}\right)^{aM}<1 .$$

When this condition is satisfied, the distance of the fast variables from the slow manifold after a macrostep can be written to lowest order as

$$|d^{n+1}|\leq\sum_{j=1}^{P}b_j\left(\frac{\lambda\Delta t}{\varepsilon}\left(1-\frac{\delta t}{\varepsilon}\right)^{aM}\right)^j|d^n| + L_hC_g(1+\lambda)\Delta t .$$

PROOF. We substitute the macrosolver (4.7) and increments (4.6) into $|d^{n+1}|$, obtaining

$$|d^{n+1}|=\left|x^{n+1}-h_0(y^{n+1})\right|$$

$$=\left|x_{M_1}^{n,1}-\frac{\lambda\Delta t}{\varepsilon}\sum_{j=1}^{P}b_jd_{M_j}^{n,j}-h_0(y^{n+1})\right|$$

$$=\left|d_{M_1}^{n,1}-\frac{\lambda\Delta t}{\varepsilon}\sum_{j=1}^{P}b_j d_{M_j}^{n,j}+h_0(y_{M_1}^{n,1})-h_0(y^{n+1})\right|$$

$$\leq\frac{\lambda\Delta t}{\varepsilon}\sum_{j=1}^{P}b_j\left|d_{M_j}^{n,j}\right|+\left|h_0(y_{M_1}^{n,1})-h_0(y^{n+1})\right|+|d_{M_1}^{n,1}|\,.$$

Employing Lemma 5.9, Assumption (*A1*) on the Lipschitz continuity of the approximate slow manifold and (4.3)–(4.5) to bound $|h_0(y_{M_1}^{n,1})-h_0(y^{n+1})|$ produces

$$|d^{n+1}|\leq\frac{\lambda\Delta t}{\varepsilon}\sum_{j=1}^{P}b_j\left(\left(1-\frac{\delta t}{\varepsilon}\right)^{M_j}|d_0^{n,j}|+L_h C_g\varepsilon\right)$$

$$+L_h\left|y_{M_1}^{n,1}-y^{n+1}\right|+\left(1-\frac{\delta t}{\varepsilon}\right)^{M_1}|d^n|+L_h C_g\varepsilon$$

$$\leq\frac{\lambda\Delta t}{\varepsilon}\sum_{j=1}^{P}b_j\left(1-\frac{\delta t}{\varepsilon}\right)^{aM}|d_0^{n,j}|+L_h C_g(1+\lambda)\Delta t$$

$$+\left(1-\frac{\delta t}{\varepsilon}\right)^{M_1}|d^n|+L_h C_g\varepsilon\,.$$

Substituting the bound from Lemma 5.17 into $|d_0^{n,j}|$ yields

$$|d^{n+1}|\leq\sum_{j=1}^{P}b_j\left(\left(\frac{\lambda\Delta t}{\varepsilon}\left(1-\frac{\delta t}{\varepsilon}\right)^{aM}\right)^{j}|d^n|+\left(L_h C_g(1+\lambda)\Delta t\right.\right.$$

$$\left.\left.+\left(1-\frac{\delta t}{\varepsilon}\right)^{M_1}|d^n|+L_h C_g\varepsilon\right)\sum_{k=0}^{j-1}\left(\frac{\lambda\Delta t}{\varepsilon}\left(1-\frac{\delta t}{\varepsilon}\right)^{aM}\right)^{k}\right)$$

$$\leq\sum_{j=1}^{P}b_j\left(\frac{\lambda\Delta t}{\varepsilon}\left(1-\frac{\delta t}{\varepsilon}\right)^{aM}\right)^{j}|d^n|+\left(L_h C_g(1+\lambda)\Delta t\right.$$

$$\left.+\left(1-\frac{\delta t}{\varepsilon}\right)^{M_1}|d^n|+L_h C_g\varepsilon\right)\sum_{j=0}^{P-1}\left(\frac{\lambda\Delta t}{\varepsilon}\left(1-\frac{\delta t}{\varepsilon}\right)^{aM}\right)^{j}\,.$$

$$(5.33)$$

The upper bound on $|d^n|$ diverges as n increases unless

$$\frac{\lambda\Delta t}{\varepsilon}\left(1-\frac{\delta t}{\varepsilon}\right)^{aM}<1\,,$$

completing the Lemma.

Remark 5.19. *The bound presented above for PI1 is not sharp. If the duration of the microsolver is sufficiently large with $M\delta t\gg\varepsilon$, or if the fast variables are*

initialised on the slow manifold with $|d^0 = 0|$, then the PI1 increments may be approximately tangent to the slow manifold and higher order accuracy in Δt can be achieved.

We now formulate analogous results for PI2.

Lemma 5.20. *Given assumptions (A1), (A3) and (A4), the distance of the fast variables from the approximate slow manifold after the n-th macrostep in the PI2 scheme, given by $|d^{n+1}| = |x^{n+1} - h_0(y^{n+1})|$, satisfies the recurrence relation*

$$|d^{n+1}| \leq \frac{\left(1 - \frac{\delta t}{\varepsilon}\right)^{aM}}{a} \left[\sum_{j=1}^{P} b_j \left(\frac{\lambda \Delta t}{\varepsilon} \left(1 - \frac{\delta t}{\varepsilon}\right)^{aM} \right)^j |d^n| \right.$$

$$\left. + L_h C_g (1 + \lambda) \Delta t \sum_{j=0}^{P-1} \left(\frac{\Delta t \lambda}{\varepsilon} \left(1 - \frac{\delta t}{\varepsilon}\right)^{aM} \right)^j \right] + 2 L_{h'} C_g^2 t_\Delta^2$$

$$+ \mathcal{O}\left(\varepsilon, \left(1 - \frac{\delta t}{\varepsilon}\right)^{M_1} |d^n| \right) .$$

In particular, the fast variables do not diverge if

$$\frac{\left(1 - \frac{\delta t}{\varepsilon}\right)^{aM}}{a} \left(\frac{\lambda \Delta t}{\varepsilon} \left(1 - \frac{\delta t}{\varepsilon}\right)^{aM} \right)^j < 1 \ \forall j.$$

When this condition is satisfied, the distance of the fast variables from the slow manifold after a macrostep can be written to lowest order as

$$|d^{n+1}| \leq \frac{\left(1 - \frac{\delta t}{\varepsilon}\right)^{aM}}{a} \sum_{j=1}^{P} b_j \left(\frac{\lambda \Delta t}{\varepsilon} \left(1 - \frac{\delta t}{\varepsilon}\right)^{aM} \right)^j |d^n| + 2 L_{h'} C_g^2 t_\Delta^2 .$$

We remark that the first bound presented in the above Lemma for PI2 is precisely $\left(1 - \frac{\delta t}{\varepsilon}\right)^{aM} / a$ times the bound presented for PI1 in Lemma 5.18, with an additional term proportional to t_Δ^2 and the curvature of the slow manifold, measured by $L_{h'}$.

PROOF. We reformulate x^{n+1}, employing (4.18) and (4.15) and Lemma 5.9 to estimate

$$|d^{n+1}| = \left| x^{n+1} - h_0(y^{n+1}) \right|$$

$$= \left| x_{M_1}^{n,1} + \sum_{j=1}^{P} \frac{b_j}{a_{j+1}} \left(x_{M_{j+1}}^{n,j+1} - x_{M_1}^{n,1} \right) - h_0(y^{n+1}) \right|$$

$$\leq \left| h_0(y_{M_1}^{n,1}) - h_0(y^{n+1}) + \sum_{j=1}^{P} \frac{b_j}{a_{j+1}} \left(h_0(y_{M_{j+1}}^{n,j+1}) - h_0(y_{M_1}^{n,1}) \right) \right|$$

$$+\sum_{j=1}^{P}\frac{b_j}{a_{j+1}}\left|d_{M_{j+1}}^{n,j+1}+(a_{j+1}-1)d_{M_1}^{n,1}\right|. \tag{5.34}$$

The first term on the right-hand side of (5.34) can be Taylor expanded to second order to obtain

$$h_0(y_{M_1}^{n,1})-h_0(y^{n+1})=h_0(y_{M_1}^{n,1})-h_0\left(y_{M_1}^{n,1}+\sum_{j=1}^{P}b_jk_{y,j}\right)$$

$$=-\mathrm{D}h_0\left(y_{M_1}^{n,1}\right)\left(\sum_{j=1}^{P}b_jk_{y,j}\right)$$

$$-\sum_{|\alpha|=2}\frac{1}{\alpha!}\partial^\alpha h_0\left(y_{M_1}^{n,1}\right)\left(\sum_{j=1}^{P}b_jk_{y,j}\right)^\alpha+\mathcal{O}(t_\Delta^3), \tag{5.35}$$

where we used multi-index notation to denote the second order derivatives of h_0. Similarly, the second term on the right-hand side of (5.34) can be estimated by Taylor expanding the chord $h_0(y_{M_{j+1}}^{n,j+1})-h_0(y_{M_1}^{n,1})$ to second order, employing (4.14), with

$$h_0(y_{M_{j+1}}^{n,j+1})-h_0(y_{M_1}^{n,1})=h_0\left(y_{M_1}^{n,1}+(y_{M_{j+1}}^{n,j+1}-y_{M_1}^{n,1})\right)-h_0(y_{M_1}^{n,1})$$

$$=\mathrm{D}h_0(y_{M_1}^{n,1})(y_{M_{j+1}}^{n,j+1}-y_{M_1}^{n,1})$$

$$+\sum_{|\alpha|=2}\frac{1}{\alpha!}\partial^\alpha h_0\left(y_{M_1}^{n,1}\right)\left(y_{M_{j+1}}^{n,j+1}-y_{M_1}^{n,1}\right)^\alpha+\mathcal{O}(t_\Delta^3)$$

$$=a_{j+1}\mathrm{D}h_0(y_{M_1}^{n,1})k_{y,j}$$

$$+a_{j+1}^2\sum_{|\alpha|=2}\frac{1}{\alpha!}\partial^\alpha h_0\left(y_{M_1}^{n,1}\right)k_{y,j}^\alpha+\mathcal{O}(t_\Delta^3). \tag{5.36}$$

Substituting (5.36) and (5.35) into (5.34) yields

$$|d^{n+1}|\leq\left|\sum_{|\alpha|=2}\frac{1}{\alpha!}\partial^\alpha h_0\left(y_{M_1}^{n,1}\right)\right|\max_{|\alpha|=2}\left|\sum_{j=1}^{P}b_ja_{j+1}k_{y,j}^\alpha-\left(\sum_{j=1}^{P}b_jk_{y,j}\right)^\alpha\right|$$

$$+\sum_{j=1}^{P}\frac{b_j}{a_{j+1}}\left|d_{M_{j+1}}^{n,j+1}+(a_{j+1}-1)d_{M_1}^{n,1}\right|+\mathcal{O}(t_\Delta^3)$$

$$\leq 2L_{h'}\max_{\substack{|\alpha|=2,\\1\leq k\leq P}}|k_{y,j}(x^n,y^n)|^\alpha$$

$$+\sum_{j=1}^{P}\frac{b_j}{a_{j+1}}\left(|d_{M_{j+1}}^{n,j+1}|+|d_{M_1}^{n,1}|\right)+\mathcal{O}(t_\Delta^3),$$

where we used that $\sum_{j=1}^{P} b_j = 1$, $a_j \leq 1$, and employed Assumption (*A3*) on the Lipshitz continuity of the Jacobian Dh_0. Employing Lemma 5.9 and recalling that the time step covered by each PI2 increment is t_Δ, we obtain

$$|d^{n+1}| \leq 2L_{h'}C_g^2 t_\Delta^2 + \sum_{j=1}^{P} \frac{b_j}{a_{j+1}} \left(\left(1 - \frac{\delta t}{\varepsilon}\right)^{aM} \left(|d_0^{n,j+1}| + |d_0^{n,1}|\right) + 2L_h C_g \varepsilon \right) ,$$

which on substituting Lemma 5.17 becomes

$$|d^{n+1}| \leq \frac{\left(1 - \frac{\delta t}{\varepsilon}\right)^{aM}}{a} \left[\sum_{j=1}^{P} b_j \left(\frac{\lambda \Delta t}{\varepsilon} \left(1 - \frac{\delta t}{\varepsilon}\right)^{aM} \right)^{j} |d^n| + \left(L_h C_g (1 + \lambda) \Delta t \right. \right.$$

$$\left. + \left(1 - \frac{\delta t}{\varepsilon}\right)^{M_1} |d^n| + L_h C_g \varepsilon \right) \sum_{j=0}^{P-1} \left(\frac{\Delta t \lambda}{\varepsilon} \left(1 - \frac{\delta t}{\varepsilon}\right)^{aM} \right)^{j} \right]$$

$$+ 2L_{h'} C_g^2 t_\Delta^2 + \mathcal{O}(\varepsilon) .$$

The upper bound on $|d^n|$ diverges as n increases unless

$$\frac{\left(1 - \frac{\delta t}{\varepsilon}\right)^{aM}}{a} \left(\frac{\lambda \Delta t}{\varepsilon} \left(1 - \frac{\delta t}{\varepsilon}\right)^{aM} \right)^{j} < 1 \; \forall j \,,$$

completing the Lemma.

Combining Lemmas 5.18 and 5.20 yields Theorem 5.3.

6. Numerics

We now illustrate the key results of Theorem 5.2 with a fourth-order Runge-Kutta macrosolver, which we recall here for $P = 4$ including the constants obtained in the proof. We employ a forward Euler microsolver with $p = 1$ unless otherwise stated.

In order to compare PI1 and PI2 at the same computational cost, we choose the number of microsteps in the PI2 method proportionally lower so that the two methods take the same number of microsteps over one macrostep, with $M_j = M$ in PI1 and $M_j = \{M, M/2, M/2, M, M\}$ in PI2. Recalling Theorem 5.2 for $P = 4$, the discretisation error $|E_d^n| = |y^n - Y(t^n)|$ is bounded in the PI1 and PI2 method by

$$|E_d^n| \leq \frac{e^{L_G t^n}}{L_G} \left\{ C_4^* t_\Delta^4 + C_2^* M_{I,II} \delta t + 2L_g \left(9 \frac{\varepsilon}{t_\Delta} + e^{-\frac{M_{I,II} \delta t}{\varepsilon}} \right) |d_{\max}^m| + 2L_g L_h C_g \varepsilon \right\} ,$$

with $M_I = M$ for PI1 and $M_{II} = M/2$ for PI2, and where for $P = 4$, $a = 1/2$. However, the distance of the fast variables from the slow manifold after a macrostep scales differently in the two methods. Recalling Theorem 5.3,

the distance of the fast variables from the approximate slow manifold $|d^n| = |x^n - h_0(y^n)|$ is bounded for stable applications of the PI1 method by

$$|d^{n+1}| \leq \sum_{j=1}^{P} b_j \left(\frac{\lambda \Delta t}{\varepsilon} \left(1 - \frac{\delta t}{\varepsilon} \right)^M \right)^j |d^n| + L_h C_g (1 + \lambda) \Delta t \ ,$$

and for stable applications of the PI2 method by

$$|d^{n+1}| \leq 2 \left(1 - \frac{\delta t}{\varepsilon} \right)^{\frac{M}{2}} \sum_{j=1}^{P} b_j \left(\frac{\lambda \Delta t}{\varepsilon} \left(1 - \frac{\delta t}{\varepsilon} \right)^{\frac{M}{2}} \right)^j |d^n| + 2 L_{h'} C_g^2 t_\Delta^2 \ .$$

We show results for the multiscale system

$$\dot{y}_\varepsilon = -x_\varepsilon y_\varepsilon - \alpha \, y_\varepsilon^2 \tag{6.1}$$

$$\dot{x}_\varepsilon = \frac{-x_\varepsilon + \sin^2(y_\varepsilon)}{\varepsilon} \ , \tag{6.2}$$

which has stable fixed point at $(0,0)$. At lowest order in ε, the associated slow reduced system is given by

$$\dot{Y} = -Y \sin^2(Y) - \alpha Y^2 \ . \tag{6.3}$$

For higher order approximations of the slow manifold and the associated coordinate transformations relating y and Y the reader is referred to the useful webtool [29] (see also [30]).

The system (6.1)-(6.2) with initial conditions $y_\varepsilon(0) > 0$ is locally Lipschitz with Lipschitz constant $L_h = 1$ and $L_g = \max(|x_\varepsilon| + 2\alpha |y_\varepsilon|)$ where the maximum is taken over the local region around the initial conditions $(x_\varepsilon(0), y_\varepsilon(0))$ under consideration. The vectorfield of the slow dynamics (6.1) is locally bounded by $C_g = \max(|x_\varepsilon y_\varepsilon| + \alpha |y_\varepsilon|^2)$, with the maximum taken over the same region. Note that the free parameter α controls the constants $C_2^* = \alpha y_\varepsilon(0)^3 (2\alpha + \sin(2y_\varepsilon(0))) + \mathcal{O}(\sin^3(y_\varepsilon(0)))$ and $C_4^* = 16\alpha^3 y_\varepsilon(0)^3 + 8\alpha^4 y_\varepsilon(0)^5 + \mathcal{O}(\sin^5(y_\varepsilon(0)))$.

We first investigate how the discretization error $|E_d^n|$ scales with the macrostep size Δt in the PI1 and PI2 methods, when all other parameters are kept fixed (except n, to fix the final time T). Our analytical result predicts that, so long as $|d_{\max}^n|$ is small and the practical assumption $\varepsilon < M\delta t$ is satisfied, results will be divided into two regimes: for $C_P^* t_\Delta^P < C_2^* M\delta t$, the bound for $|E_d^n|$ is dominated by the term proportional to $C_2^* M\delta t$ and $|E_d^n|$ is independant of t_Δ; for $C_P^* t_\Delta^P > C_2^* M\delta t$, the scaling is $|E_d^n| \sim t_\Delta^P$. The slight advantage of the PI2 method in this case is that distributing the same total number of microsteps over more applications of the microsolver results in lower error due to $M\delta t$. To keep the term proportional to $|d_{\max}^n|$ small in both cases, we choose parameters so that $\frac{\lambda \Delta t}{\varepsilon} \left(1 - \frac{\delta t}{\varepsilon} \right)^{aM} < 1$. The predicted regimes are clearly visible in Figure 3, where results are presented for a range of macrostep sizes Δt for PI1 and

PI2. We choose $\alpha = 0.2$, and the scale separation parameter $\varepsilon = 10^{-9}$. We use $M = 40$ microsteps with microstep size $\delta t = 0.4\varepsilon$, while the number of iterations n vary from 20 to 10^5 to keep $T = 1$ fixed for all values of Δt. Initial conditions are chosen to lie on the approximate slow manifold with $y^0 = 1$, $x^0 = \sin^2(1)$. The Lipschitz constants are $L_g = 1.1$ and $L_h = 1$, the bound on the vector field of the slow dynamics is $C_g = 2$, and the maximal derivatives of the reduced slow dynamics are $C_2^* = 4$ and $C_4^* = 8$.

We present results for the error scaling of $|E_d^n|$ with the microstep size δt in Figure 4. To focus on the scaling with $M\delta t$, we select parameters with $C_P^* t_\Delta^P < C_2^* M\delta t$, and control the distance of the fast variables from the slow manifold $|d_{\max}^n|$ by ensuring $\frac{\lambda\Delta t}{\varepsilon}\left(1 - \frac{\delta t}{\varepsilon}\right)^{aM} < 1$. Figure 4 confirms our analytical result, that under the condition $C_P^* t_\Delta^P < C_2^* M\delta t$, the discretization error scales like $E_d^n \sim M\delta t$. The advantage of the PI2 method here is that distributing the microsteps over an additional application of the microsolver leads to an overall smaller error compared to PI1 due to the smaller drift of the slow variables over the microsolver.

We use a second-order Runge-Kutta microsolver (i.e. $p = 2$), to demonstrate that the scaling with δt is not affected by the order of the microsolver. We choose $\alpha = 1$, and the scale separation parameter $\varepsilon = 10^{-5}$. We use $M = 100$ microsteps and $n = 50$ iterations of each method. The macrostep size Δt varies from 0.0035 to 0.0032 to keep t_Δ fixed as δt increases. Initial conditions are chosen to lie on the approximate slow manifold with $y^0 = 5$, $x^0 = \sin^2(5)$. The Lipschitz constants are $L_g = 11$ and $L_h = 1$, the bound on the vector field of the slow dynamics is $C_g = 30$, and the maximal derivatives of the reduced slow dynamics are $C_2^* = 50$ and $C_4^* = 2000$.

We illustrate the linear scaling of $|E_d^n|$ with the maximal distance $|d_{\max}^n|$ of the fast variable from the approximate slow manifold after a macrostep in Figure 5. We do so by scaling the initial condition for the fast variables, x^0. To ensure that the error is not dominated by the initial initialization error $|d^0|$, we choose parameters which render the scheme unstable, allowing for divergence of the fast variables from the slow manifold over the macrosteps, i.e. $|d^n| > |d^0|$. Figure 5 confirms clearly the linear dependence of $|E_d^n|$ on $|d_{\max}^n|$. We choose $\alpha = 1$, and the scale separation parameter $\varepsilon = 10^{-4}$. We use $M = 100$ microsteps with microstep size $\delta t = 0.01\varepsilon$, and $n = 5$ iterations of each method with macrostep size $\Delta t = 10^{-3}$. Initial conditions are $y^0 = 1$, $x^0 \in [\sin^2(1) + 0.01, \sin^2(1) + 1]$. The Lipschitz constants are $L_g = 3$ and $L_h = 1$, the bound on the vector field of the slow dynamics is $C_g = 2$, and the maximal derivatives of the reduced slow dynamics are $C_2^* = 3$ and $C_4^* = 24$.

We investigate how $|d_{\max}^n|$, the maximal deviation of the fast variables from the slow manifold, scales with Δt in the PI1 and PI2 methods. We choose parameters satisfying $\frac{\lambda\Delta t}{\varepsilon}\left(1 - \frac{\delta t}{\varepsilon}\right)^{aM} \ll 1$, so that the bounds presented for $|d^n|$

in Theorem 5.3 imply

$$|d_{\max}^n| \le L_h C_g (1+\lambda)\Delta t + \mathcal{O}(\Delta t^2)$$

for the PI1 method, and

$$|d_{\max}^n| \le 2L_{h'} C_g^2 t_\Delta^2 + \mathcal{O}\left(\Delta t^3, \frac{(1-\frac{\delta t}{\varepsilon})^{aM}}{a}\Delta t\right)$$

for the PI2 method. As noted in Remark 5.19, the bound for the PI1 method is not tight for systems with $|d^0|=0$. We therefore choose initial conditions off the slow manifold. Furthermore, to ensure that the initial error $|d^0|$ does not dominate the error $|d_{\max}^n|$, we record $|d_{\max}^n|$ after the first macrostep. Figure 6 clearly shows the linear dependence of $|d_{\max}^n|$ with the macrostep size Δt for PI1, and the quadratic dependence of $|d_{\max}^n|$ with the macrostep t_Δ for PI2. We choose again $\alpha = 0.2$, and the scale separation parameter $\varepsilon = 10^{-9}$. We use $M=40$ microsteps with microstep size $\delta t = 0.4\varepsilon$, while the number of iterations n vary from 20 to 10^5 to keep $T=1$ fixed for all values of Δt. Initial conditions are $y^0=1$, $x^0 = \sin^2(1)+1$. The Lipschitz constants are $L_g=5$ and $L_h=1$, the bound on the vector field of the slow dynamics is $C_g=2$, and the maximal derivatives of the reduced slow dynamics are $C_2^*=4$ and $C_4^*=8$.

Finally, we investigate how $\Delta y^{T,\Delta t} = |y^{T,\Delta t} - y^{T,\Delta t/2}|$ scales with the macrostep size Δt where $y^{T,\Delta t}$ are the outputs of the PI1 or PI2 methods with macrostep size Δt and final time T. In [4] $\Delta y^{T,\Delta t}$ was used as a measure of the numerical error. In Figure 7 we show how $\Delta y^{T,\Delta t}$ scales with Δt for a fourth-order Runge-Kutta macrosolver, employing the same numerical parameters as in Figure 3. It is seen that $|y^{T,\Delta t} - y^{T,\Delta t/2}| \sim t_\Delta^4$ for all values of Δt whereas the actual discretization error is dominated by $M\delta t$ at the lower values of Δt (cf. Figure 3). Hence, such proxies for the numerical error have to be treated with caution when evaluating PI methods.

We comment that the numerical results presented here are robust; in particular, we confirm that identical scalings can be produced from simulations of the Michaelis-Menten system employed in [22], although its fast dynamics does not satisfy our form (2.4), and the Brusselator with rapidly replenished source employed in [2], where the approximate slow manifold is constant.

7. Discussion

We have introduced PI2, a seamless numerical multiscale method with a higher-order macrosolver, which is a slight modification of a standard implementation of a projective integration method, PI1, involving an additional application of the microsolver. In both PI1 and PI2, each increment is rooted on the slow manifold. In PI1 the increments typically do not end on the slow manifold. In contrast, the additional application of the microsolver assures that in PI2 each increment also ends on the slow manifold, even for slow manifolds

with non-vanishing curvature (see Figures 1 and 2). If the slow manifold is sufficiently linear over the course of one macrostep, the increments of PI2 then all lie approximately tangential to it.

We presented error bounds for the slow variables for both methods, expressed in Theorem 5.2. The error bounds are not affected by the order of the microsolver used (though strictly speaking, we only considered explicit microsolver schemes). Hence the contribution of the microsolver to the error constitutes a bottleneck for PI methods, after which the error in the slow variables cannot be improved by adjusting the macrostep size or the order of the macro- or microsolver. Hence there is no gain to be expected in the slow dynamics when microsolvers other than forward Euler schemes are used.

In Theorem 5.3 we derived bounds for the unphysical deviation of the fast variables from the slow manifold, which may cause numerical instability [12], and provided a stability criterion for the macrostep size.

The Theorems now allow us to compare the PI1 and PI2 schemes. A fair comparison requires that both schemes are operated at the same computational cost. Hence, PI2 utilises less microsteps per application of the microsolver during the construction of the increments as the total number of microsteps is distributed over one more application of the microsolver. Consequently, the absolute discretisation error of PI2 is smaller when compared to PI1. Theorem 5.3 establishes that the PI2 method incurs less deviation from the slow manifold as the deviations scale quadratically with the macrostep size rather than linearly as for PI1. The improved stability can be attributed to the increments of PI2 pointing towards the slow manifold, enforced by the additional relaxation towards the slow manifold when constructing the increments.

Acknowledgments

References

[1] I. G. Kevrekidis, C. W. Gear, J. M. Hyman, G. K. Panagiotis, O. Runborg, C. Theodoropoulos, Equation-free, coarse-grained multiscale computation: Enabling microscopic simulators to perform system-level analysis, Comm. Math. Sci. 1 (4) (2003) 715–762.

[2] C. Gear, I. Kevrekidis, Projective methods for stiff differential equations: Problems with gaps in their eigenvalue spectrum, SIAM J. Sci. Comp. 24 (4) (2003) 1091–1106.

[3] G. Hummer, I. Kevrekidis, Coarse molecular dynamics of a peptide fragment: Free energy, kinetics, and long-time dynamics computations, Journal of Chemical Physics 118 (23) (2003) 10762–10773.

[4] S. L. Lee, C. W. Gear, Second-order accurate projective integrators for multiscale problems, Journal of Computational and Applied Mathematics 201 (1) (2007) 258–274.

[5] C. Vandekerckhove, D. Roose, K. Lust, Numerical stability analysis of an acceleration scheme for step size constrained time integrators, Journal of Computational and Applied Mathematics 200 (2) (2007) 761–777.

[6] D. Givon, I. G. Kevrekidis, R. Kupferman, Strong convergence of projective integration schemes for singularly perturbed stochastic differential systems, Comm. Math. Sci. 4 (4) (2006) 707–729.

[7] I. Kevrekidis, G. Samaey, Equation-free multiscale computation: algorithms and applications, Ann. Rev. Phys. Chem. 60 (2009) 321–344.

[8] P. Lafitte, G. Samaey, Asymptotic-preserving projective integration schemes for kinetic equations in the diffusion limit, SIAM J. Sci. Comput. 34 (2) (2010) A579–A602.

[9] W. E, B. Engquist, The heterogeneous multiscale methods, Comm. Math. Sci. 1 (1) (2003) 87–132.

[10] E. Vanden-Eijnden, Numerical techniques for multi-scale dynamical systems with stochastic effects, Comm. Math. Sci. 1 (2) (2003) 385–391.

[11] W. E, Analysis of the heterogeneous multiscale method for ordinary differential equations, Comm. Math. Sci. 1 (3) (2003) 423–436.

[12] B. Engquist, Y.-H. Tsai, Heterogeneous multiscale methods for stiff ordinary differential equations, Mathematics of Computation 74 (252) (2005) 1707–1742.

[13] W. E, D. Liu, E. Vanden-Eijnden, Analysis of multiscale methods for stochastic differential equations, Communications on Pure and Applied Mathematics 58 (11) (2005) 1544–1585.

[14] W. E, B. Engquist, X. Li, W. Ren, E. Vanden-Eijnden, Heterogeneous multiscale methods: A review, Comm. Comp. Phys. 2 (3) (2007) 367–450.

[15] D. Liu, Analysis of multiscale methods for stochastic dynamical systems with multiple time scales, SIAM Multiscale Model. Simul. 8 (3) (2010) 944–964.

[16] E. Vanden-Eijnden, On HMM-like integrators and projective integration methods for systems with multiple time scales, Comm. Math. Sci 5 (2) (2007) 495–505.

[17] W. E, E. Vanden-Eijnden, Some critical issues for the "Equation-Free" approach to multiscale modeling, arXiv:0806.1621v1 [math.NA].

[18] J. Maclean, G. A. Gottwald, On convergence of the projective integration method for stiff ordinary differential equations, Comm. Math. Sci. 12 (2) (2014) 235–255.

[19] C. Gear, I. G. Kevrekidis, C. Theodoropoulos, 'coarse' integration/bifurcation analysis via microscopic simulators: micro-galerkin methods, Computers and Chemical Engineering 26 (2002) 941–963.

[20] C. Vandekerckhove, D. Roose, Accuracy analysis of acceleration schemes for stiff multiscale problems, Journal of Computational and Applied Mathematics 211 (2) (2008) 181–200.

[21] P. Lafitte, A. Lejon, G. Samaey, A high-order asymptotic-preserving scheme for kinetic equations using projective integration, preprint, arXiv:1404.6104v2.

[22] C. Gear, T. J. Kaper, I. G. Kevrekidis, A. Zagaris, Projecting to a slow manifold: Singularly perturbed systems and legacy codes, SIAM J. Appl. Dyn. Syst. 4 (3) (2005) 711–732.

[23] A. Zagaris, C. Gear, T. Kaper, I. Kevrekidis, Analysis of the accuracy and convergence of equation-free projection to a slow manifold, Math. Mod. Num. Anal. 43 (4) (2009) 757–784.

[24] C. Vandekerckhove, B. Sonday, A. Makeev, D. Roose, I. Kevrekidis, A common approach to the computation of coarse-scale steady states and to consistent initialization on a slow manifold., Computers & Chemical Engineering 35 (10) (2011) 1949–1958.

[25] A. Zagaris, C. Vandekerckhove, C. Gear, T. Kaper, I. Kevrekidis, Stability and stabilization of the constrained runs schemes for equation-free projection to a slow manifold., Discrete and Continuous Dynamical Systems - Series A 32 (8) (2012) 2759–2803.

[26] C. I. Siettos, C. W. Gear, I. G. Kevrekidis, An equation-free approach to agent-based computation: Bifurcation analysis and control of stationary states, EPL (Europhysics Letters) 99 (4) (2012) 48007.

[27] A. Iserles, A First Course in the Numerical Analysis of Differential Equations, Cambridge University Press, Cambridge, 2009.

[28] J. Carr, Applications of Centre Manifold Theory, no. 35 in Applied Mathematical Sciences, Springer, 1981.

[29] A. J. Roberts, Slow manifold of stochastic or deterministic multiscale differential equations, http://www.maths.adelaide.edu.au/anthony.roberts/sdesm.php (2008).

[30] A. J. Roberts, Normal form transforms separate slow and fast modes in stochastic dynamical systems, Physica A 387 (1) (2008) 12–38.
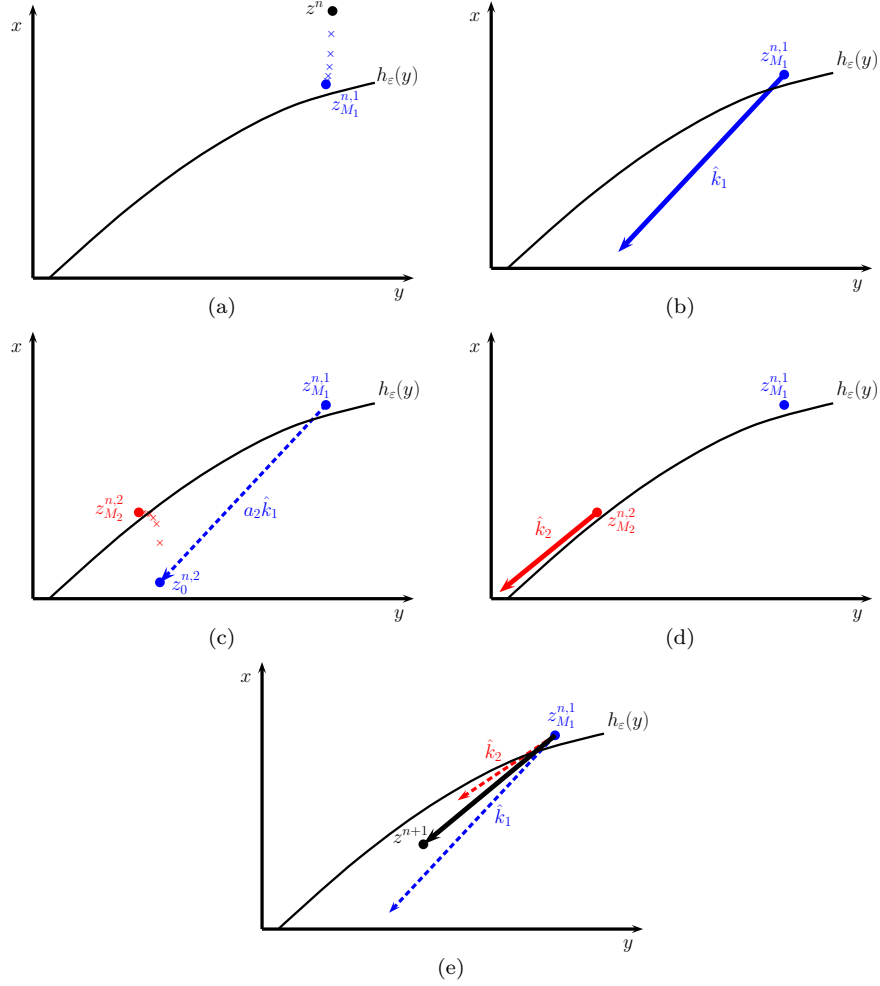
Figure 1: Sketch of the PI1 scheme for a second order Runge-Kutta macrosolver. The microsolver is employed in 1a, 1c, the increments $\hat{k}_j$ are given by vectorfield evaluations in 1b, 1d, and the macrosolver is illustrated in 1e. For this scheme $a_1 = 0$, $a_2 = 1$ and $b_1 = b_2 = 1/2$.
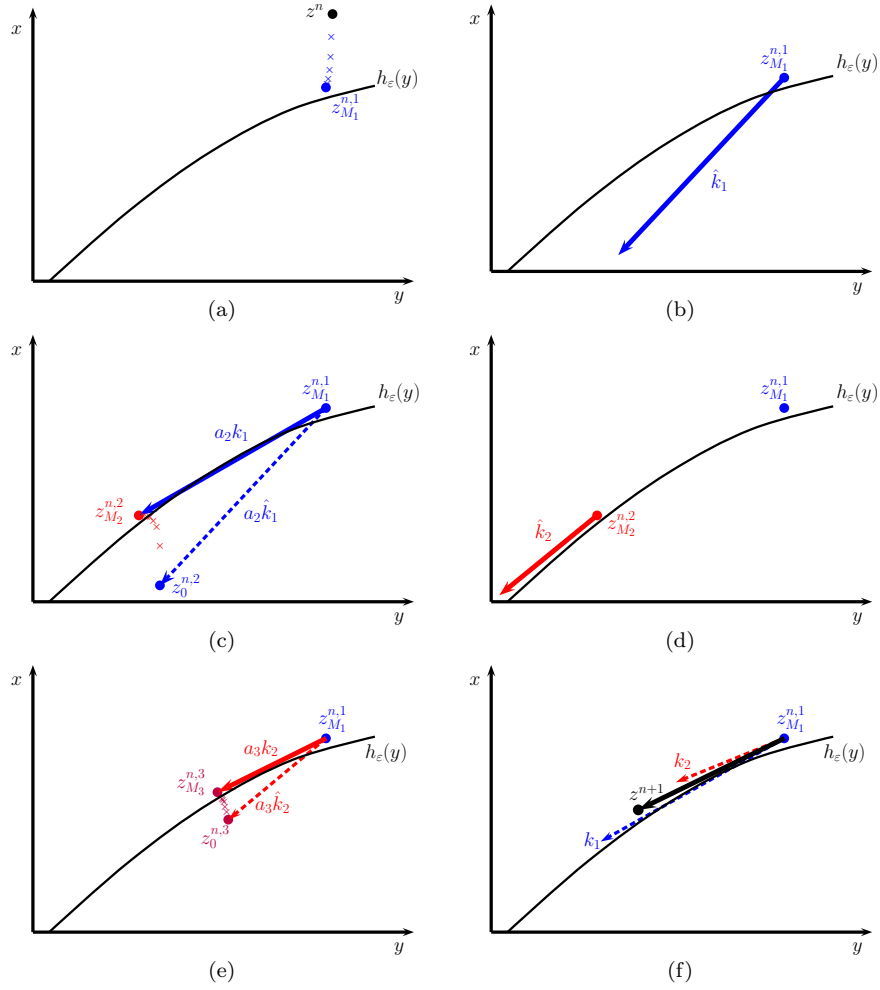
Figure 2: Sketch of the PI2 scheme for a second order Runge-Kutta macrosolver. The microsolver is employed in 2a, 2c, 2e and the (now auxiliary) quantities $\hat{k}_j$ are given by vector field evaluations in 2b, 2d. The increments k_j are estimated in 2c, 2e, and the macrosolver is illustrated in 2f. For this scheme $a_1 = 0$, $a_2 = a_3 = 1$ and $b_1 = b_2 = 1/2$.
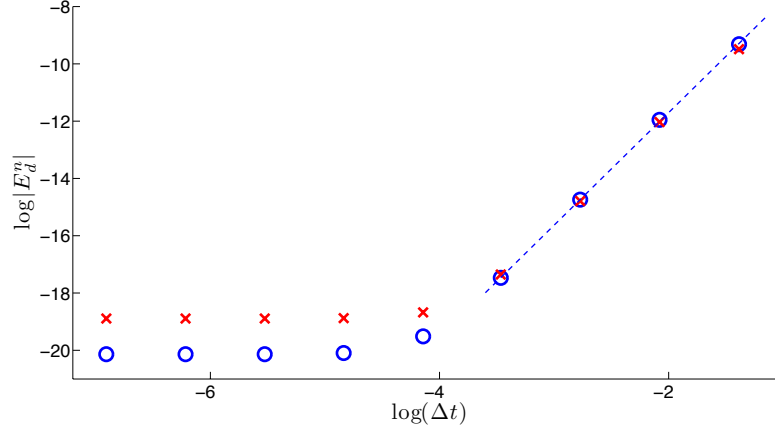
Figure 3: Plot of $\log|E_d^n|$ versus $\log(\Delta t)$ for fixed time of integration $T=1$ of the system (6.1)-(6.2). The crosses represent results from the PI1 scheme and the circles represent results from the PI2 scheme. The dashed line is a linear regression line with a slope of 3.93.
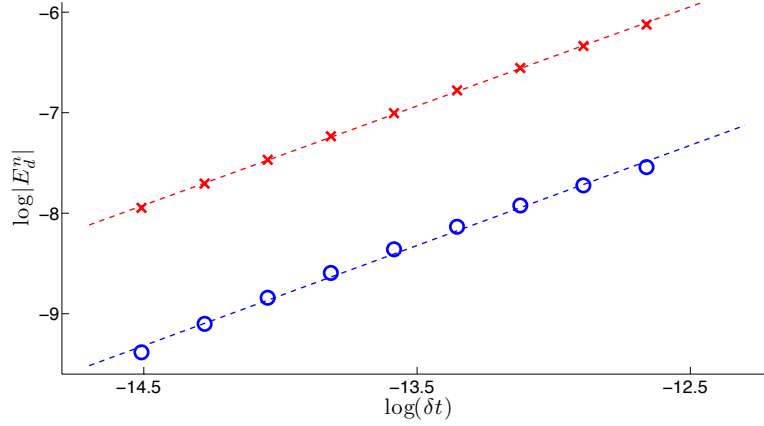

Figure 4: Plot of $\log|E_d^n|$ versus $\log(\delta t)$ for fixed time of integration $T=0.18$ of the system (6.1)-(6.2). The crosses represent results from the PI1 scheme and the circles represent results from the PI2 scheme. The dashed lines are linear regression lines with a slopes of 0.99 and 1.0, respectively.
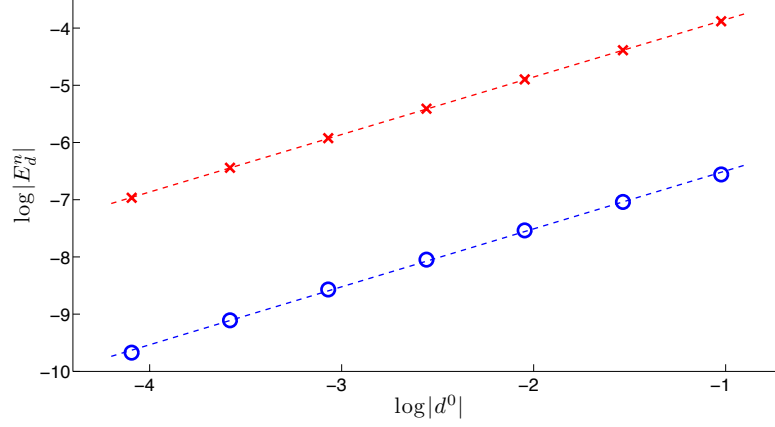
Figure 5: Plot of $\log|E_d^n|$ versus $\log|d^0|$ for fixed time of integration $T = 0.0056$ of the system (6.1)-(6.2). The crosses represent results from the PI1 scheme and the circles represent results from the PI2 scheme. The dashed lines are linear regression lines with slopes of 1.01.
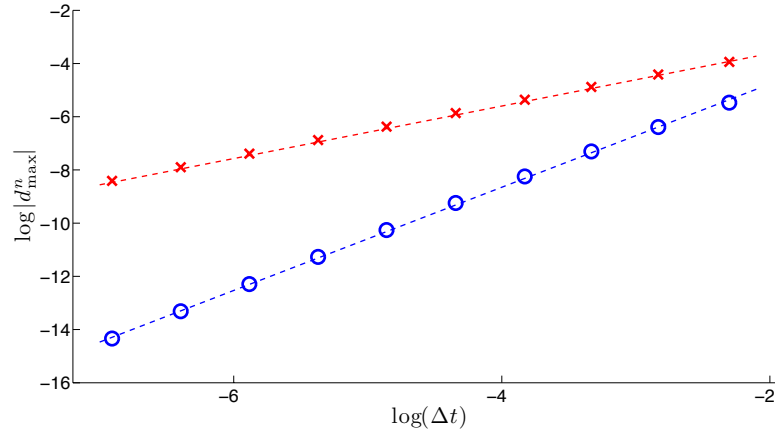

Figure 6: Plot of $\log|d_{\max}^n|$ versus $\log(\Delta t)$ for fixed time of integration $T = 1$ of the system (6.1)-(6.2). The crosses represent results from the PI1 scheme and the circles represent results from the PI2 scheme. The dashed lines are linear regression lines with a slopes of 0.98 and 1.94, respectively.
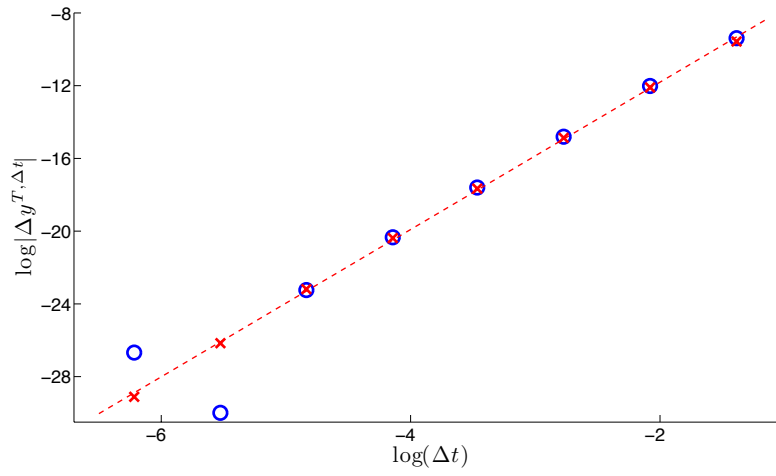
Figure 7: Plot of $\log|\Delta y^{T,\Delta t}|$ versus $\log|\Delta t|$ for fixed time of integration $T=1$ of the system (6.1)-(6.2). The crosses represent results from the PI1 scheme and the circles represent results from the PI2 scheme. The dashed line is a linear regression line (of the PI1 output) with slope of 4.05.